J Clin Epidemiol Vol. 44, No. 6, pp. 483–496, 1991
Printed in Great Britain.

0895-4356/91 $3.00 + 0.00

A RANDOMIZED CLINICAL TRIAL OF NEGATIVE PRESSURE VENTILATION IN SEVERE CHRONIC OBSTRUCTIVE PULMONARY DISEASE: DESIGN AND METHODS

STANLEY H. SHAPIRO,[1]* PETER T. MACKLEM,[2,3] KATHERINE GRAY-DONALD,[1,4] JAMES G. MARTIN,[2,3] PIERRE P. ERNST,[1] SHARON WOOD-DAUPHINEE,[1,5] TOM A. HUTCHINSON[1,3] and WALTER O. SPITZER[1]

Departments of [1]Epidemiology and Biostatistics and [2]Medicine, McGill University, [3]Montreal Chest Hospital Centre, [4]School of Dietetics and Human Nutrition, McGill University, and [5]School of Physical and Occupational Therapy, McGill University, Montreal, Quebec, Canada

(Received in revised form 9 November 1990)

Abstract—This report documents the design and methods of a randomized clinical trial designed to test the effectiveness of home negative pressure ventilation in patients with severe chronic obstructive pulmonary disease. Active negative pressure ventilation was compared with a sham version of the treatment after a pre-trial assessment had indicated the feasibility of the latter. Over 1200 patients in the metropolitan Montreal area were screened. Of these, 348 patients were recruited to enter a 4-week stabilization period, and 184 were subsequently randomized to receive either active or sham negative pressure ventilation. A 5-day in-hospital period was used to train patients in ventilator use and obtain baseline measures of exercise capacity, lung function, respiratory symptoms, and quality of life. Home ventilation treatment took place during a following 12-week period. Respirator use was recorded both from patient logs and from concealed meters installed in the units. Patients received four home visits by physiotherapists during the 12-week period and returned for follow-up to the hospital 4 and 12 weeks post-discharge for reassessment.

Chronic obstructive pulmonary disease Randomized clinical trial Negative pressure ventilation

1. INTRODUCTION

In the U.S. the total amount of disability payments to patients with chronic obstructive pulmonary disease (COPD) is second only to disability payments to patients with coronary disease. In Canada the annual morbidity from COPD results in an estimated 50,000 hospitalizations involving half a million hospital bed-days [1]. Data from the U.S. Health Interview Survey [2] suggest: (1) that 1 to 1.5 million persons in the U.S.A. take medicine or some other form of treatment recommended by a physician for chronic bronchitis/emphysema; (2) that nearly 1 million patients see a physician at least five times a year for these conditions; and (3) that more than half of these individuals have at least 14 days of disability in bed each year (emphysema alone accounts for an estimated 19,000,000 person-days in bed annually in the U.S.A.) [3]. Combined economic costs of chronic obstructive lung disease for

*All correspondence should be addressed to: Dr S. Shapiro, Department of Epidemiology and Biostatistics, McGill University, 1020 Pine Avenue West, Montreal, Quebec, Canada.

health care, time lost from work, and lost wages have been estimated to be as high as U.S. $15 billion per year [3]. A cheap effective means of rehabilitation would be of major benefit.

This report documents the design and methods of a randomized clinical trial of the potential effectiveness of one such intervention, negative pressure ventilation (NPV). The trial was designed to address the question of whether negative pressure ventilation confers benefit to patients with severe COPD by putting muscles to rest and represents an effective treatment for such patients. The nature of the intervention posed methodologic challenges exceeding those typically arising in a trial of a pharmacologic or surgical intervention. Given the complex study logistics and the attendant problems of an appropriate control, this article focuses on the trial's design and methods. The results of the trial will be reported in a separate future article.

Much of the disability accompanying COPD results from patients' sensation of dyspnea, or shortness of breath, which limits their ability to exercise. Although agreement is not uniform, most chest physicians and respiratory physiologists attribute this sensation to afferent signals from the inspiratory muscles or to their drive from the respiratory centres. It is plausible that in advanced disease dyspnea is aggravated by inspiratory muscle fatigue. Because these patients must breathe continually against an excessive load, they may never be able to obtain the rest required for their muscles to recover by themselves. If so, mechanical ventilation might lead to recovery from fatigue, a lessening of dyspnea and an improved exercise performance.

The role of respiratory muscle rest in the rehabilitation of patients with severe COPD was supported in a preliminary study by Braun and Marino [4] who reported substantial improvement in exercise performance and arterial blood gas tensions in case studies of patients with COPD who underwent nocturnal negative pressure ventilation. Prompted by this finding and other work on inspiratory muscle fatigue [5–8], we planned a randomized controlled clinical trial to determine if nocturnal negative pressure ventilation accomplished with home use of a ventilator improves exercise performance, alleviates the sensation of dyspnea and improves quality of life of patients with severe COPD.

2. MAJOR FEATURES OF THE TRIAL

2.1. Intervention

In the past, mechanical ventilation has been accomplished using cumbersome devices such as the iron lung or various techniques such as endotrachial intubation. Newer methods are much more manageable and apply negative pressure to the body surface with a chest wall cover such as a suit sealed at the patient's hips, arms and neck. We employed a respirator pump made by Puritan Bennett. At the time of the study it was the quietest and hence the most conducive to sleep. The pneumo-suit, manufactured by Nu-Tech Corp., is held away from the chest wall by a grid which fits over the chest and upper abdomen. A widebore tube connected to a fitting in the suit above the grid transmits negative pressure to the chest wall from the pump thereby inflating the lungs. The apparatus is pictured in Fig. 1.

2.2. Pre-trial assessment of sham treatment

We were concerned that improvement in patients receiving NPV could be a combination of the direct physiologic effects of NPV as well as the indirect effects of being intervened upon. Placebo and Hawthorne effects [9] are well recognized problems in the evaluation of interventions. When there is no standard therapy the use of a placebo control in drug trials has become almost a *sine qua non*. However, the use of a placebo for the evaluation of medical devices is considerably more problematic.

We anticipated that sham NPV, in which the treatment would be applied with only minimal negative pressure, would be associated with adverse side effects. In particular, discomfort due to restricted movement and possible sleep disruption of the patients or spouse due to the noise of the respirator were expected to be problems. We initially believed that this would offset any favourable placebo effects and that there would be little likelihood of patients remaining both compliant and blinded to the specific form of NPV that they were receiving. However, we decided to examine this assumption in a pilot study and to pretest the logistics of patient enrollment, treatment, and outcome assessment. We made such assessments with 34 patients enrolled before the period of the definitive trial and randomized on a 2 to 1 basis to sham or active NPV.

Several steps were taken to attempt to assure that patients remained "blinded" during the

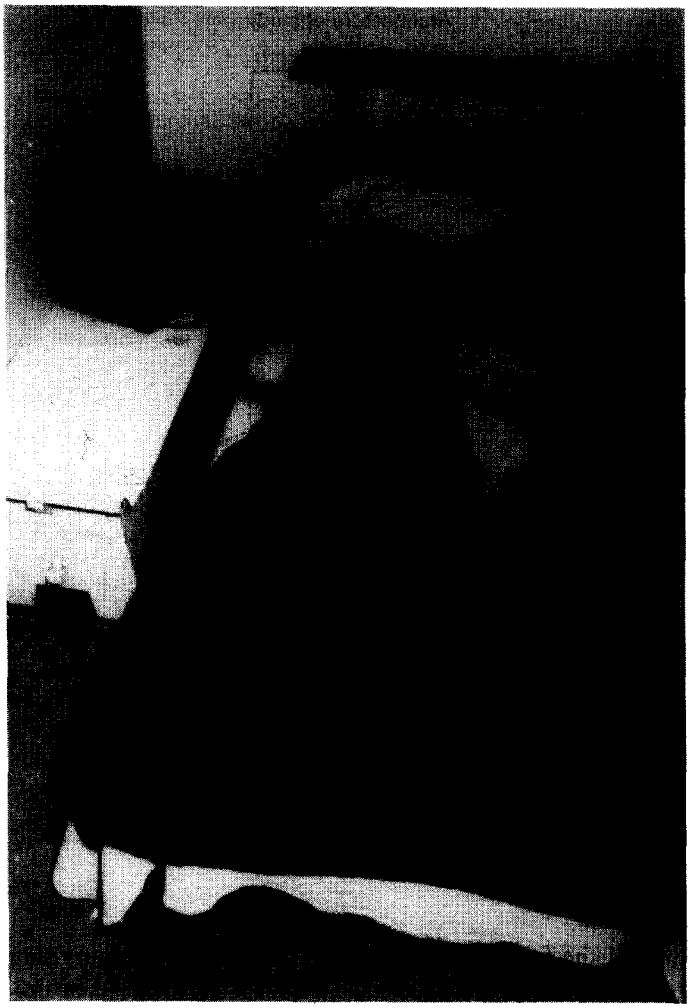

Fig. 1. Apparatus employed for NPV.

pre-trial. Patients were randomized to either sham NPV or active NPV after the first day of a 5-day period of in-hospital training, and irrespective of their assigned group, underwent the same training and testing programs. As in the active group, patients undergoing sham treatment spent two sessions in the physiology laboratory where measures of diaphragmatic activity were made to select the ventilatory parameters for the respirator. This was done to preserve blinding as patients allocated to sham treatment arbitrarily had the negative pressure set at a minimal level, typically -5 cm H_2O. This setting was sufficient to cause some motion of the body suit similar to that seen with active treatment, but not sufficient to provide ventilatory muscle rest. The pressure settings used were not revealed to the patients. Arterial oxygen saturation was monitored for all patients during the hospital stay.

To our surprise, we found that sham treatment was tolerated as well as active NPV, and that with our procedures patients did not in fact become unblinded to the form of ventilation they were receiving. The pilot study also showed that nocturnal NPV was not well suited to a substantial proportion of subjects, irrespective of treatment status. Only 52% (11/21) of the sham group and 31% (4/13) of the active treatment group completed the entire 3-month program. The most frequent problem reported by patients was difficulty in sleeping. It appeared that for many patients inability to tolerate nocturnal NPV was a major cause of discouragement. For this reason in the definitive trial we decided to allow patients the option of receiving NPV during the day.

Given our preliminary results, we decided to employ a sham NPV arm in the full trial. Our pilot studies thus led to the final design, a two-arm randomized clinical trial of active vs

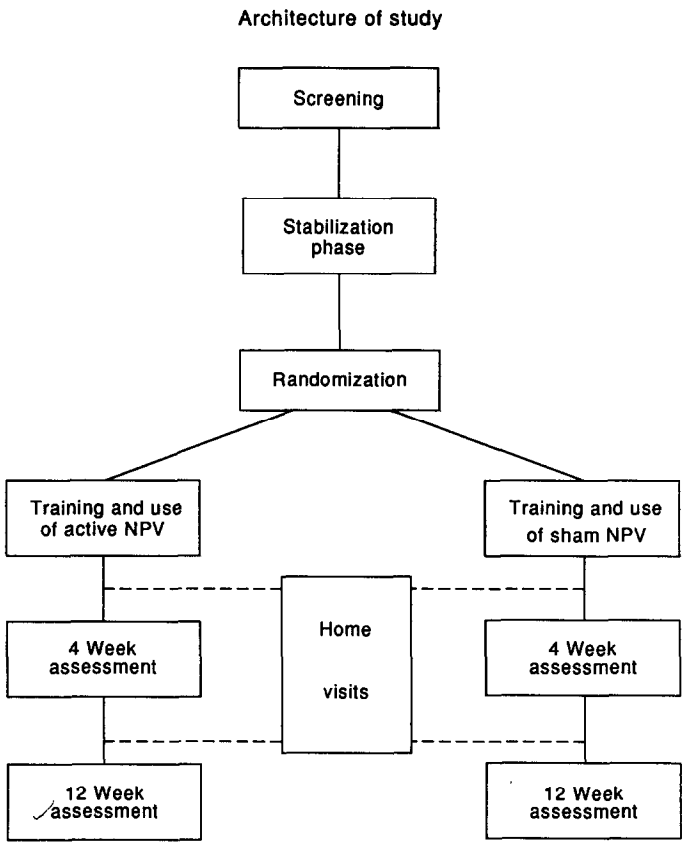

Fig. 2

sham NPV. An outline of the final architecture adopted for the study is provided in Fig. 2.

2.3. Primary outcome measure

Dyspnea was an obvious candidate for the primary study outcome given its central role in the disability experienced by patients with severe COPD. However, in common with many other symptoms, it is difficult to assess with precision. Despite recent developments [10, 11] we were reluctant to use the level of dyspnea or its change over time as primary outcome variables.

We felt that a clinically significant improvement in dyspnea would be reflected in improved exercise tolerance, which could be more accurately quantified. Recent research [12–14] points to a close link between dyspnea, exercise performance and the ventilatory muscles. When subjects who have severe COPD exercise, dyspnea usually supervenes and limits the maximal performance. This level of exercise often corresponds to the maximal predicted ventilatory capacity of the subject and suggests that ventilatory muscle function may be an important determinant of exercise capacity in these subjects.

Thus, at the planning stage, we decided to use exercise tolerance, measured by the endurance time on a cycle ergometer set at 2/3 of the peak power output established on a symptom limited incremental exercise test [15] as the primary outcome variable. A second measure of exercise tolerance was the 6-minute walk test [16] that provided the maximum distance a subject walks in 6 minutes in a measured corridor. Other outcomes of interest included respiratory symptoms, resting lung function, arterial blood gases, and quality of life. These are discussed in more detail in the subsequent section on baseline measurements.

During the feasibility phase we experienced difficulties in assessing endurance times in the way that we had initially envisioned. The standard deviation of the change in endurance times was 8.47 and 4.72 minutes in the sham and treatment group respectively in comparison with the estimate of 2.5 minutes that had come from the literature [17]. This difficulty persisted into the actual trial itself and led to a re-evaluation of the suitability of endurance time as the primary outcome. After consideration of the variability of the outcome measures by the

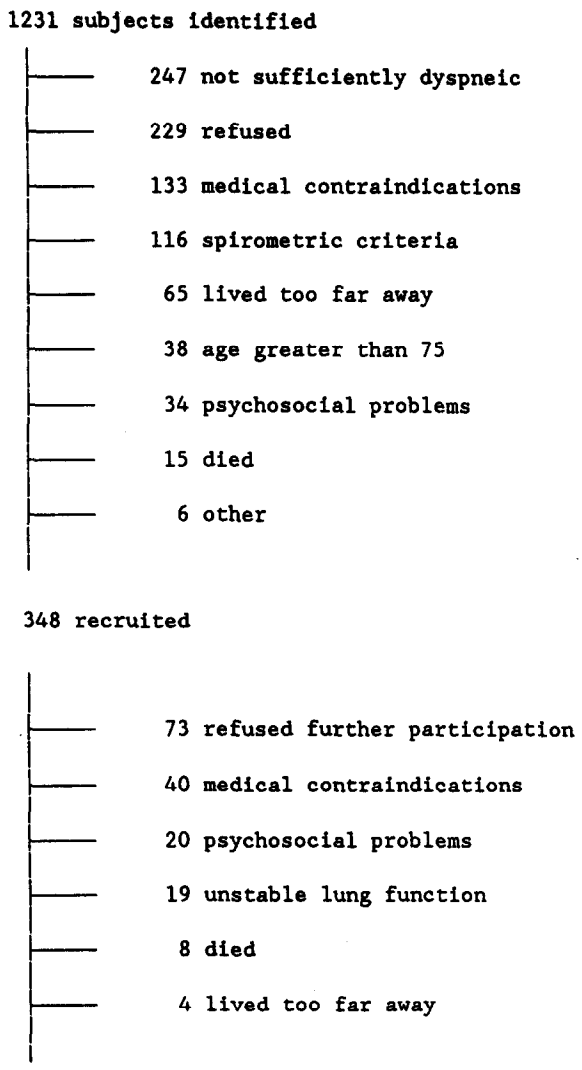

Fig. 3

Steering Committee of the trial (whose voting members were totally unaware of differences in the level of the variable between treatment groups) it was decided to shift to the 6-minute walk as the primary measure of exercise tolerance. This decision was endorsed by the independent Safety and Data Monitoring Committee established by the National Heart, Lung, and Blood Institute (NHLBI) to monitor the progress of the trial.

3. PATIENT SELECTION AND ASSIGNMENT

3.1. Eligibility criteria

Patients with severe symptomatic chronic obstructive lung disease were candidates for entry into the study. Eligibility criteria for the trial are listed in Table 1. The criteria fall into several categories reflecting both the severity and stability of the underlying disease as well as suitability to fulfil study requirements. Grading of the severity of dyspnea proceeded according to the American Thoracic Society [18] modification of the Medical Research Council respiratory questionnaire. A dyspnea level of grade 4 reflected the fact that the patient felt they had to stop for breath after walking 100 yards on the level; a level of grade 5 indicated that the patient felt too breathless to leave the house, or became breathless when dressing or undressing. The spirometric criteria were chosen to be similar to those used in other randomized controlled clinical trials in populations with COPD. If there was concern about the adequacy of patients' mental status a Pfeiffer mental status assessment was carried out [19].

3.2. Recruitment

Initial recruitment activities were concentrated at the Montreal Chest Hospital Centre and the Home Care Programme of Maisonneuve-Rosemont Hospital, a special respiratory home care programme. As a result of our experience during the pre-trial we discovered that recruitment efforts needed to be greatly expanded. Screening to search for eligible patients was done in 15 major hospitals in the Montreal area. In collaboration with treating physicians, charts were initially reviewed to identify patients with COPD who were between 30 and 75 years of age, whose most recent FEV_1 (forced expiratory volume in 1 second) was not greater than 60% predicted, and whose chart review revealed no obvious medical contraindications. A letter describing the study was then sent to these patients and followed up by a telephone call to determine their interest in the study and their availability for further screening.

In addition, personal letters were sent to every practising respiratory physician in the Montreal area, and newspaper, television and radio coverage were arranged to solicit volunteers. The distribution of the sources of referral to the trial is shown in the first column of Table 2. In some settings the patients were better known to the referring physician, particularly in the Maisonneuve-Rosemont Home Care Programme, and patients not thought to be ultimately eligible were not referred. In contrast, self-referred patients had not been pre-evaluated. As a consequence, the proportion of randomized subjects from each source varied widely as seen in the right-hand column of Table 2.

Table 1. Eligibility criteria

Eligibility for Entry into Study
—A clinical diagnosis of COPD
—Age between 30–75 years
—Dyspnea grade 4 or 5 as assessed on the ATS symptom questionnaire and not known to be due to a condition other than COPD
—Pre-bronchodilator $FEV_1 < 50\%$ predicted
—Pre-bronchodilator $FEV_1/FVC < 0.6$
—Post bronchodilator $FEV_1 < 60\%$
—Ability to walk and exercise on a cycle ergometer
—Residence close enough to the Montreal Chest Hospital to be accessible for home and clinic visits
—Fluency in English or French
—No apparent risk of acute medical emergencies or conditions which would effect patients' participation such as severe unstable angina, history of acute pulmonary edema, morbid obesity, alcoholism
—Adequate mental ability (as determined by Pfeiffer mental status assessment if necessary)
—Use of salbutamol inhaler and either one or both of ipratropium bromide inhaler and oral theophylline
—Post-stabilization FEV_1 within 25% of recruitment FEV_1
—Post-stabilization dyspnea score not more than one grade higher than recruitment level
—No exacerbation during stabilization period

A total of 1231 subjects were interviewed. Approximately 1/4 of these, 348, met the initial eligibility criteria and agreed to participate in the study. Reasons for non-participation of the 883 other patients are listed in Table 3. The most frequent reasons were lack of sufficient dyspnea, patient refusal, and medical contradictions.

3.3. Stabilization

The 348 patients recruited participated in a 4-week stabilization period during which time they were visited at home by a nurse who provided the patient and family member with further information and assessed the candidate's suitability in terms of living arrangements and social support. The nurse also verified that the patient was taking at least two bronchodilators

Table 2. Source of patients identified and randomized

Hospitals	Number identified	Number randomized
Montreal Chest Hospital	130	15
Maisonneuve-Rosemont Home Care Program	51	26
Montreal General Hospital	58	7
Hôpital St Luc	19	9
Hôpital Hotel Dieu	33	9
Hôpital Notre Dame	67	10
Hôpital Sacré Coeur	35	4
Verdun General Hospital	103	11
Hôpital Cité de la Santé	45	7
Other hospitals	161	13
Self referral (via media)	529	71
	1231	184

Table 3. Reasons for non-participation of screened patients

1231	*Patients identified*
247	Not sufficiently dyspneic
229	Refused
133	Medical contraindications
116	Spirometric criteria
65	Lived too far away
38	Age greater than 75
34	Psychosocial problems
15	Died
6	Other
348	*Patients recruited*
73	Refused further participation
40	Medical contraindications
20	Psychosocial problems
19	Unstable lung function
8	Died
4	Lived too far away
184	*Patients randomized*

on a regular basis, salbutamol inhaler and either one or both of ipratropium bromide inhaler and oral theophylline. If the subject was taking a theophylline preparation and if the serum theophylline level at recruitment was not between 40 and 110 mmol/l the dosage was altered and a blood sample drawn approximately 1 week later to verify that adequate serum concentration had been obtained.

Three to four weeks after recruitment, patients were re-evaluated to assess lung function stability and to detect any unreported medical contraindications. During stabilization an additional 164 patients were excluded. Of these, 73 declined further participation, 40 had new or unreported medical contraindications, 20 had psychological or social problems precluding participation, 19 had unstable lung function or sensation of dyspnea (6 improved on their dyspnea rating, 6 had unstable spirometry and 7 had exacerbations of their condition), 8 patients died during this period, and 4 patients could not be included owing to the distance from home to hospital.

3.4. Randomization

Patients with stable lung function were eligible to enter hospital for instruction in NPV and to undergo baseline testing. Randomization to active or sham NPV took place on the second day of hospitalization after a final evaluation designed to eliminate patients who were unable to perform the required tests and training or in whom testing seemed contraindicated. We randomized eligible patients according to a stratified block randomization scheme. Stratification was based on whether or not the patient

had been receiving home oxygen. We felt that home oxygen status provided a composite measure of many different baseline factors and as such was an important variable upon which to achieve treatment balance. Randomization proceeded within strata according to a permuted block scheme with a block size, or balancing interval, varying randomly between 8 or 12 according to the outcome of a computer generated random number [20]. This ensured that the cumulative number of assignments to each treatment would be in balance after each block of assignments had been made. Treatment assignments were obtained by telephone request to the trial secretary who was located away from the site of clinical activities.

4. BASELINE MEASUREMENTS

The distribution of baseline characteristics is provided in Table 4. All measures were carried out by staff blinded as to treatment group.

4.1. Respiratory symptoms

The American Thoracic Society respiratory symptom questionnaire [18] was administered by a trained interviewer. A more detailed evaluation of the symptom of dyspnea was made with two additional instruments. The first, a recently developed clinical questionnaire [10] allows breathlessness to be rated for each of three separate features; magnitude of task which elicits the symptom, magnitude of effort required to elicit the symptom, and the resultant functional impairment. This instrument was designed as an unstructured interview to be used by the clinician as part of the usual history taking. In an attempt to reduce bias and improve reproducibility, a structured questionnaire was developed and subsequently administered by a trained interviewer. The scoring for this questionnaire was described by the original authors [11]. Up to four points are given for each of the three components, magnitude of task, effort, and functional impairment, with a score of 12 indicating "no dyspnea" despite strenuous effort as well as "no limitation" of any activities due to breathlessness. Breathlessness was further quantified using an oxygen-cost diagram as described by McGavin and coworkers [21]. In the diagram everyday activities are listed and placed proportionately to their oxygen cost along a vertical line and subjects are asked to mark with a pen the point above which they think their breathlessness limits their activity.

4.2. Quality of life

Global quality of life was measured by the Quality of Life Index developed by Spitzer and colleagues [22]. The Index contains five items reflecting day-to-day activity patterns, self-care capabilities, general health, outlook on life and support by family and friends. Each item is scored as 0, 1 or 2 and scores are summed providing an unweighed total score (0–10).

The Index has demonstrated acceptable internal consistency [22–24], and good interrater reliability [22, 25] and stability [26, 27]. In validation studies, it has been able to distinguish between healthy and sick individuals [22] and can differentiate patients in various stages of

Table 4. Baseline characteristics of randomized patients

		Active		Sham	
	n	Mean	(SD)	Mean	(SD)
Age (years)	184	63.6	(7.32)	65.2	(7.97)
% Male	184	70%		75%	
% on home O_2	184	13%		14%	
FEV_1 (% pred)	184	29%	(9.1%)	30.0%	(11.3%)
FEV_1/FVC	184	0.33	(0.09)	0.32	(0.07)
P_aO_2 (mmHg)*	184	70.9	(9.36)	71.3	(11.83)
P_aCO_2 (mmHg)	184	43.8	(6.66)	44.4	(7.30)
PI_{max} (cmH_2O)	182	41.7	(14.72)	44.4	(15.72)
PE_{max} (cmH_2O)	182	84.4	(40.51)	82.0	(32.57)
DL_{co}	172	13.0	(4.83)	12.9	(4.69)
Body mass index (kg/m^2)	184	23.6	(4.89)	24.0	(4.36)
% ATS grade 5	184	86%		83%	
6-minute walk (m)	184	319	(115.2)	315	(103.3)
Cycle ergometer (min)	181	7.84	(6.56)	6.86	(4.37)
Oxygen cost score (mm)	184	51.3	(15.38)	49.3	(13.94)
Quality of life	184	6.5	(2.01)	6.7	(1.76)

*Measurements for patients receiving home O_2 were taken while patients received supplemental oxygen.

a disease [23]. It correlates well with other measures of quality of life [22, 23, 25, 27], is able to predict survival [28] and is responsive to clinical changes over time [29]. The Index can be completed in less than 1 minute by patients, their relatives or health care professionals in English or French. Although it has not been used as an endpoint in a clinical study of patients with chest disease, it was designed for use with both cancer patients and those with other chronic debilitating conditions. It has been extensively employed in studies of cancer and other medical and surgical conditions [23, 27, 28, 30, 32].

4.3. Anthropometric measures

Anthropometric measures including height, weight, triceps skinfold (TSF), and mid-arm circumference (MAC) were obtained by a single technician. The site of the measurement for the TSF and MAC was located at the midpoint between the acromion and the olecranon process. Three measurements were taken and the mean of the closest two readings was used [33].

4.4. Resting lung function

Spirometry was done with a nose-clip and one Vitalograph spirometer calibrated daily using a 1.5 liter calibrated syringe. The usual 3-liter syringe was not used given our interest in the greatest precision in the low range of volumes recorded in our study subjects. Volumes were corrected to BTPS (body temperature and pressure saturated). The best FEV_1 and FVC (forced vital capacity) were selected from two reproducible volume–time curves as recommended by the American Thoracic Society [23].

Single breath diffusing capacity for carbon monoxide (DLCO) was measured using a Morgan gas transfer system (P. K. Morgan Ltd, Chatham, Kent, England). The average of the two highest values was selected if these were within 10% or if not achieved in five trials we selected the highest DLCO measured corrected to standard hemoglobin concentration of 14.6 g dl^{-1} [27]. The same standard test gas was used throughout the study. The zero point was set and the system checked for leaks before each test. Biological calibration was carried out weekly by having the technician carry out the test herself.

Maximal respiratory mouth pressures were carried out in a body plethysmograph using a vacuumed 1002 mouthpiece attached to ± 200 cm H_2O validyne pressure transducer. Maximal inspiratory pressure was measured at functional residual capacity providing a small air leak via a 14 gauge needle through the mouth-piece. Maximal expiratory pressure (PE_{max}) was measured at total lung capacity. The best of three trials was reported for each of maximal inspiratory (PI_{max}) and PE_{max}. The pressure transducer was calibrated weekly with a water manometer. Instrumentation laboratory blood gas analysers (Fisher Scientific, Lenenpton, Mass.) were adjusted and calibrated daily using artificial blood controls from the same company.

4.5. Exercise capacity and endurance

The principal outcome measure of this randomized controlled trial was the distance in meters walked in 6 mintues. The method is based on the 12-minute walking test described by McGavin [34]. We chose to use a 6-minute test because it yields comparable results to the 12-minute test [16], and subjects already had a prolonged testing schedule to complete at baseline. Patients receiving home oxygen walked with oxygen delivered by nasal prongs at the rate they usually used for exercise. Subjects pulled their own portable oxygen tank on wheels. They were instructed to walk back and forth along a hospital corridor 20 meters long and to cover as much distance as possible during the 6 minutes. An encouraging phrase [35] was given by the tester each time the subject returned to the point of origin. The subject was warned when there was only 1 minute left. Because improvement in performance subsequent to familiarization with this test has been demonstrated previously [34, 35] a practice walk-test was done one day prior to the baseline measure. At least 30 minutes of rest was arranged between the 6-minute walk and the cycle endurance test.

On the first day of the hospital stay subjects underwent a symptom-limited maximal incremental exercise test on an electronically braked cycle ergometer (Bosch). Subjects breathed through a mouthpiece attached to a low resistance valve with a combined deadspace of 200 ml. Because of worsening dyspnea with such added deadspace, subjects newly enrolled in the second half of the study underwent cycle ergometer testing with a tight-fitting mask (Hans Rudolf No. 7900 m). For subjects on home oxygen an equivalent inspired oxygen concentration was provided using a 220 liter Douglas bag. After ventilation and heart rate had been

stable for at least 1 minute, subjects began cycling at approximately 50 rpm at 0 workload. The electronically determined workload was increased to 10 then 20 W at 1 and 2 minutes, then increased by 5 W at each minute thereafter. Maximum power output was defined as the level of work maintained for 30 seconds or more. The following day, after randomization, subjects returned to the exercise laboratory where they were asked to cycle to a symptom-limited maximum at a constant workload equal to two-thirds the maximal workload reached the previous day. The duration of cycling at this workload was taken as the cycle endurance time.

Variability in cycle endurance time in the first 6 months of the trial was clearly greater than expected, and two factors responsible at least in part for this excess variability were identified. Firstly, subjects who were able to cycle for 20 minutes or more had probably not achieved maximal exercise capacity on the progressive test. They were therefore exercising for long periods at relatively low workloads during the endurance test and were often stopping because of musculo-skeletal pain or even on occasion boredom. Secondly, it became apparent that subjects cycling at workloads less than 25 W had less reproducible endurance tests at the follow-up visits. This was due to a substantial electronic overestimation of the actual workload levels below 25 W.

To remedy this situation the following changes in the exercise protocol were introduced prospectively. When selecting the load for endurance testing a minimal load of 25 W was chosen, unless the subject had not maintained this load for a least 20 seconds during the incremental test. If this was the case the endurance test was done at 20 W but at a constant pedal speed of 50 rpm. If a subject's endurance time at baseline reached 20 minutes, the test was stopped and repeated the next day at a workload increased by 10 W. The endurance time at the higher workload was used as the baseline measurement.

5. TRAINING IN-HOSPITAL

5.1. Ventilator use

During the 5-day in-hospital period patients were trained in the use of the respirator, the equipment was adapted to the individual, necessary baseline evaluations were carried out as described above, and additional orientation and instruction were provided for patients and their families regarding the study.

On the first day, patients were shown a videotape which provided general information about the study and about the respirator they would be using. Patients were also given the opportunity to try on the pneumo-suit and practice getting into it independently. The respirator itself was not used so that baseline data, including oximetry, could be obtained during the initial day and night of hospitalization. On the morning of the second hsopital day the final decision was made regarding patient eligibility. Patients were randomized to either active or sham treatment and then transferred to the physiology laboratory. There NPV was instituted by a technician, supervised by a study physician, and measurement of the degree of suppression of electrical activity of the diaphragm was made.

During this first session in the physiology laboratory, appropriate ventilator settings were established. Patients subsequently received several practice sessions with the NPV equipment on the second and third hospital days. Although the equipment limits movement, adequate ventilation can be obtained in supine, lateral or semi-recumbent (in a reclining chair) positions. Patients were encouraged to choose the most comfortable position for NPV since sessions optimally were to last for 4 or more hours and periods of use of less than 2 hours duration were strongly discouraged. On the third day of hospitalization the physical therapist who was to do the home care follow-up also visited the patient. At this time she made arrangements for the first home visit, became acquainted with the patient, and established plans for installation of the equipment in the home. The following day both active and sham treatment patients had their second session in the physiology laboratory. At this time we verified the degree of phasic diaphragmatic EMG suppression with the predetermined NPV settings. Since some patients chose new positions for improved comfort we also adjusted the ventilatory settings at this time. Patients also used the day for additional practice with the respirator and to solve problems dealing with ventilator use that had surfaced during the week.

On the fifth day patients were discharged home and the equipment was delivered by the hospital therapist. Patients were reminded how to complete a logbook to record respirator usage. At home they installed themselves in

the suit and the whole apparatus was checked. The hospital physical therapist also left a brief summary of the patient's history for the visiting therapist and all the forms necessary for each follow-up visit.

5.2. Capture

Through the use of negative pressure ventilation we sought to rest the respiratory muscles. We assessed the extent to which this was achieved by measuring the degree of suppression of electrical activity of the diaphragm, or "capture". All measurements were performed on both sham and active treatment groups first while the patient was resting comfortably in the suit and breathing normally, then with the ventilator in operation. These measurements were: (a) pressure within the suit; (b) thoraco-abdominal movement as assessed by inductance plethysmography; (c) electrical activity of the diaphragm as assessed by the peak value of the integrated electromyogram (EMG) signal obtained from surface electrodes placed on the 6th and 7th intercostal spaces near the anterior axillary line; in a preliminary study we showed that this was highly correlated with the electrical activity of the diaphragm recorded from an esophageal electrode; (d) arterial oxygen saturation and pulse rate as obtained by finger oximetry. Signals were recorded on an eight-channel strip chart recorder.

After at least 5 minutes of quiet breathing as defined by the regularity of thoraco-abdominal motion, steadiness of the size of the integrated EMG signal, and the constancy of the arterial oxygen saturation, the ventilator was turned on. Patients assigned to sham ventilation received the least amount of negative pressure generated with the ventilator on; this usually ranged between -4 and -8 cm H_2O. These patients received a ventilatory rate equivalent to their spontaneous respiratory rate. For patients assigned to active treatment we administered a negative pressure and ventilatory rate consistent with the aim of maximizing patient comfort and at the same time reducing the integrated EMG by at least 50% and suppressing visible spontaneous respiratory effort.

The second capture session was carried out in the position that the patient had indicated as being their position of preference. Patients were sent home with the settings established in the second capture session. The mean negative pressure setting for active patients was -30 cm H_2O with 75% of the settings between -28 cm H_2O and -32 cm H_2O. At their first follow-up visit the capture criteria were re-evaluated and ventilatory settings changed if necessary. Table 5 presents the distribution of percent reduction in EMG from the last capture session for those patients assigned to active treatment. Overall, the median reduction achieved was 65%. A random sample of 20 sham patients had their EMG tracings read to estimate their percent reduction. The mean (SD) reduction in this group was $10 \pm 12.5\%$.

Table 5. Suppression of integrated diaphragmatic EMG at last capture session in active negative pressure group ($n = 88$)*

Percent suppression	Percentage of patients
<50	18%
50–60	18%
60–70	22%
70–80	23%
>80	19%

*Two patients left hospital before capture could be assessed and in two others repeated attempts to obtain usable electromyogram tracing were unsuccessful.

6. FOLLOW-UP PROCEDURES

The home follow-up regimen alternated follow-up visits and phone calls. Each patient routinely received four follow-up visits by the same therapist. Each follow-up visit followed a standard plan. Patients demonstrated the ability to get into the machine and to turn it on. The therapist discussed how the machine was being used, the time spent in the machine on a daily basis and per session and any reason for non-use of the machine. The pressures and rates were checked by the therapist as was the clock which recorded usage. Phone calls to the patient were conducted by the same visiting therapist. She inquired about any problem areas, encouraged the patients, and provided general support.

Patients were scheduled to return to the Montreal Chest Hospital at 4 weeks and 12 weeks into the study to meet with their study physician and to perform required tests. Originally, we planned to have an 8-week evaluation as well, but found that this burden of assessment was poorly tolerated by patients. The cycle endurance tests were particularly anxiety provoking for most patients.

We therefore eliminated assessment of outcome at 8 weeks, although each patient was contacted by a study physician at that time. At the hospital follow-up visits, the questions on dyspnea were repeated. A change in dyspnea was also assessed from a transition score for

the clinical questionnaire used at baseline [10]. Briefly a transition score of −3 (major deterioration) to +3 (major improvement) is given for a change in each of the three components of the questionnaire, magnitude of task, effort and functional impairment. As for the baseline assessment, a questionnaire was developed so that it could be administered by a trained interviewer. The oxygen-cost diagram was initially re-administered in an analogous fashion to baseline. Subsequently patients were also asked to mark on the same form their recollection of their baseline level. This provided us with the opportunity to examine two measures of change. Assessment of quality of life was carried out in the same manner as the baseline quality of life assessment.

At each follow-up visit spirometry was repeated using the same spirometer. Mouth pressures and arterial blood gases were also repeated. Both the 6-minute walk and the cycle endurance time were repeated on the same day at 1 and 3 months and in the same order in which they were done at baseline. If the patient was receiving oxygen, the same concentration in inspired air was always used. All changes in the cycle exercise protocols were introduced prospectively so that patients exercised at follow-up according to the same protocol that had been used at baseline for their initial evaluation. Patients receiving oxygen breathed a fixed concentration of 28% O_2 for the cycle endurance test. Height, weight, skinfold thickness and arm circumference were re-measured at 3 months.

7. STATISTICAL CONSIDERATIONS

7.1. Analysis plan

Given the dangers of focusing analysis on subgroups of patients defined by their post-randomization behaviour, primary data analysis will follow an "intention to treat" strategy and include all randomized individuals, irrespective of level of respirator usage or follow-up experience. The last available follow-up measurement will be used to assess change. Those without a follow-up assessment will be arbitrarily assigned a change of O. Secondary analysis examining selected groups of patients will be undertaken to supplement the primary examination. This will include analyses restricted to those patients actually providing follow-up, as well as analyses of subgroups defined by their baseline characteristics including patients who were hypercapniac, those who were receiving home oxygen, those who have predicted FEV_1 values less than 25%, and those whose cycle endurance tests were performed at workload below 25 W.

For the primary outcome of 6-minute walk distance, the comparison of the treatment and sham groups will be carried out using multiple linear regression. Differences between the treatment groups will be assessed by regressing the final 6-minute walk distance on the baseline walk, home oxygen status, FEV_1% pred, and P_aCO_2. Analysis of secondary outcomes will proceed in an analogous fashion. In addition, the effect of other potential correlates of outcome, including hours of respirator use, will be explored.

7.2. Sample size

Our preliminary planning had been based on the assumption that the principal outcome would be the change in cycle endurance time, and that the primary comparison of interest would be the difference in the mean change in endurance times between the active and sham groups. Based on an anticipated standard deviation of 4.5 minutes, a minimum sample size of 168 patients was targeted in order to achieve a power of 0.90 to detect a true difference of 2.25 minutes between the experimental and control groups [36]. As previously discussed the subsequent difficulty in obtaining reliable estimates of the change in endurance times led to the decision to use the 6-minute walk test as the principal outcome. Given the greater relative precision of the 6-minute walk test, this change did not require any increase in the minimum targeted sample size; with a minimum sample size of 168 patients we would have a power in excess of 0.90 of detecting true differences between groups as small as 25 meters.

8. ASSESSMENT OF BLINDING AND COMPLIANCE

All procedures were identical for both active and sham groups including both in-hospital training and assessment, and home follow-up visits. Patients had private rooms in hospital, and training programs were individualized. Although treating personnel were aware of patients' status, those involved in the assessment of outcome were not. Treating personnel were all employed by the study and were aware that patients and outcome assessors were to remain blinded. Two separate systems of patient record storage were established, with potentially unblinding data not available to blinded personnel.

Table 6. Guessed treatment

Actual treatment	Sham	Active	Could not say	
Sham	26	47	11	84
Active	7	65	6	78
	33	112	17	162

Weekly meetings of treating personnel were held to discuss any problems of patient management. At the end of the study patients were asked to indicate the treatment they believed they had been assigned. One hundred and sixty-two patients provided an answer and their responses, classified according to their actual treatment group, are provided in Table 6. Of the 84 sham patients, 47, or 56%, guessed that they had been receiving active treatment; the comparable figure for the active treatment groups was 65 out of 78 or 83%.

Of the 184 patients randomized, 10 patients did not use the respirators at their residence at all. Of these 10 patients, 2 were sham and 8 were active. An additional 53 patients, 27 sham and 26 active, received only partial treatment in that they stopped treatment prior to completing 12 weeks. The most frequent reasons given by patients for treatment dropout are given in Table 7.

For all patients respirator usage was assessed in two ways. Unknown to the patients a clock was installed in the respirator to record the time of use. Patients were also instructed to record their use of the machines in a daily log which was subdivided into hourly intervals. This provides us with data on the pattern, as well as total hours of use of the machine. An independent reading of total usage was obtained from the meter which was installed in the machine.

Table 7. Reasons for non-compliance by treatment group

	Active	Sham
No use of respirator		
Hospitalized	2	0
Choking sensation	1	0
Refusal	2	1
Family problems	3	0
Alzheimers disease	0	1
Total	8	2
Some use of respirator		
Patient died	2	1
Hospitalized	9	8
Choking sensation	1	0
Chest pain or tightness	3	0
Other discomfort (backpain, arthritis, etc.)	5	3
Tired, depressed, intervention too difficult	4	12
Family problems	2	0
Accidental injury	0	2
Total	26	27

9. DISCUSSION

Conceptually, this experimental trial has been straightforward. What is unusual and virtually unique is the use of a "placebo" for a therapeutic intervention as logistically and technically complicated as NPV. The major challenges have been logistic ones in the field. We succeeded in giving treatment to patients in the control group with a sham pneumo-suit and ventilator, maintaining blindness of patients as to their treatment assignment. This adds a new dimension to placebo control of clinical trials.

In addition to measurement of exercise capacity and endurance, we measured respiratory function, carried out a detailed evaluation of dyspnea, assessed quality of life (QL) and obtained anthropometric, clinical and pertinent paraclinical data. All these measurements and tests on people who were quite sick required a great effort of coordination and much diligence on the part of research support staff as well as patient cooperation.

An exhaustive search for eligible patients in the entire metropolitan Montreal area was carried out. The extensive recruitment efforts suggest to us that the majority of potentially eligible patients with chronic obstructive lung disease in a geographically defined area were identified and evaluated as part of the trial activities.

As it was our hypothesis that respiratory muscle rest would improve dyspnea and exercise performance by treating chronic inspiratory muscle fatigue, our emphasis in establishing eligibility criteria was focused on dyspnea and abnormal lung mechanics rather than on abnormalities in gas exchange. Thus CO_2 retention was not an admission criterion. This distinguishes our study from several published uncontrolled studies [4, 37, 38]. Despite this fact, almost a third of our patients, 56, were hypercapnic. This number is greater than the total sample size of any of the above-mentioned studies, and we will readily be able to examine the effect of NPV in this subgroup.

Results from uncontrolled studies or poorly controlled studies have generated a phase of enthusiasm for negative pressure ventilation in the management of patients with advanced COPD. Accordingly, the results of

this randomized controlled trial, whether they are positive or negative, will be clinically important and provide valid evidence for physicians seeking cost-effective treatment for their patients with this very difficult and disabling disease. Given the high prevalence of the disease in the population, the high burden of suffering it creates for society and for families, and the adverse economic impact of the disability, the economic implications of either positive or negative findings will be important in the formulation of health policy for the care of people with chronic obstructive lung disease.

Acknowledgements—This research work was supported by a grant from the National Heart Lung and Blood Institute, NIH. Dr James Martin was supported in part by a Medical Research Council of Canada Research Scientist Award, Dr Pierre Ernst was a recipient of a Fraser-Monat-McPherson Faculty Award from McGill University. We would like to acknowledge the assistance of Drs Robert Levy, James Hanley, Sharyn Mannix, Bernard Chrysler, Zev Carrey and Alain Beaupré in the planning and conduct of the study. We also wish to thank Marie-Annik Barbier, Katherine Berg, Carl Brewer, Gail Clements, Patricia Furci, Laurie Gibbons, Diane Laforte, Sheila Moriarty, Dan Nguyen, Fay Nurse, Ann Savoie and Sheila Segal for their efforts.

REFERENCES

1. Statistics Canada. **Hosp Morbidity Canadian Diagnostic List,** 1975.
2. **Prevalence of Selected Chronic Respiratory Conditions.** Rainbow Series No. 84. U.S. Vital and Health Statistics; 1975.
3. Public Services Laboratory, Georgetown University. Costs of illness, fiscal year, 1975. **National Heart Lung and Blood Fact Book.** DHEW Publ. No. NIH 79-1656; 1978.
4. Braun N, Marino WD. Effect of daily intermittent rest of respiratory muscles in patients with severe chronic airflow limitation (abstr). **Chest** 1984; 85: 595S.
5. Roussos CS, Macklem PT. Diaphragmatic fatigue in man. **J Appl Physiol Respir Environ Exercise Physiol** 1977; 43(2): 189–197.
6. Macklem PT. The respiratory muscles: vital pump. **Chest** 1980; 78: 753–758.
7. Macklem PT. Hyperinflation. **Am Rev Resp Dis** 1984; 129: 1–2.
8. Rabinovitch B, Pardy RL, Hussain SNA, Macklem PT. The acute effects of rest on ventilatory muscle (VM) function in patients with severe chronic airflow limitation. **Physiologist** 1983; 26: A-21.
9. Kramer MS, Shapiro SH. Scientific challenges in the application of randomized trials. **J Clin Med Assoc** 1984; 252: 2739–2745.
10. Stoller JK, Ferranti R, Feinstein AR. Further specification and evaluation of a new clinical index for dyspnea. **Am Rev Resp Dis** 1986; 134: 1129–1134.
11. Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The measurement of dyspnea. **Chest** 1984; 85: 751–758.
12. Ward ME, Eidelman D, Stubbing DG, Bellemare F, Macklem PT. Respiratory sensation and pattern of respiratory muscle activation during diaphragm fatigue. **Am Physiol Soc** 1988; 2181–2189.
13. Bradley TD, Chartrand DA, Fitting JW, Killian KJ, Grassino A. The relation of inspiratory effort sensation to fatiguing patterns of the diaphragm. **Am Rev Resp Dis** 1986; 134: 1119–1124.
14. Stubbing DG, Ramsdale EH, Killian KJ, Campbell EJM. Psychophysics of inspiratory muscle force. **J Appl Physiol** 1983; 54: 1216–1221.
15. Jones NL, Campbell MD. **Clinical Exercise Testing.** Philadelphia, Penn: W. B. Saunders; 1982.
16. Butland RJA, Pang J, Gross ER, Woodcock AA, Geddes EM. Two-, six-, and twelve-minute walking test in respiratory disease. **Br Med J** 1982; 284: 100–108.
17. Pardy *et al.* Inspiratory muscle training compared with physiotherapy in patients with chronic airflow limitation. **Am Rev Resp Dis** 1981; 123: 421–425.
18. Ferris BG. Epidemiology standardization project. **Am Rev Resp Dis** 1978; 118: 1–119.
19. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. **Am Geriatr Soc** 1975; 23: 433–441.
20. Friedman LM, Finberg CD, De Mets DL. **Fundamentals of Clinical Trials.** Littleton: PSG Publishing; 1985.
21. McGavin CR, Artvinli M, Naoe H, McHardy GJR. Dyspnoea, disability, and distance walked: comparison of estimates of exercise performance in respiratory disease. **Br Med J** 1978; 2: 241–243.
22. Spitzer WO, Dobson AJ, Hall J, Chesterman E, Levi J, Shepherd R, Battista RN, Catchlove BR. Measuring the quality of life of cancer patients. A concise QL-Index for use by physicians. **J Chron Dis** 1981; 34: 585–597.
23. Mor V. Cancer patients' quality of life over the disease course: lessons from the real world. **J Chron Dis** 1987; 40: 535–544.
24. Mor V, Stalker MZ, Gralla R, Scher HI, Cimma C, Park D, Flaherty AM, Kiss M, Nelson P, Laliberte K, Schwartz R, Marks PA, Oettgen HF. Day hosital as an alternative to inpatient care for cancer patients: a random assignment trial. **J Clin Epidemiol** 1988; 41: 771–785.
25. Gough IR, Furnival CM, Schilder L, Grove W. Assessment of the quality of life of patients with advanced cancer. **Eur J Cancer Clin Oncol** 1983; 19: 1161–1165.
26. Slevin ML, Plant H, Lynch D, Drinkwater J, Gregory WM. Who should measure quality of life, the doctor of the patient? **Br J Cancer** 1988; 57: 109–112.
27. Levine MN, Guyatt GH, Gent M, De Pauw S, Goodyear MD, Hryniuk WM, Arnold A, Findlay B, Skillings JR, Bramwell VH, Levin L, Bush H, Abu-Zahra H, Kotalik J. Quality of life in stage II breast cancer: an instrument for clinical trials. **J Clin Oncol** 1988; 6: 1798–1810.
28. Coates A, Gebski V, Bishop JF, Jeal PN, Woods RL, Snyder R, Tattersall MHN, Byrne M, Harvey V, Gill G, Simpson J, Drummond R, Browne J, Van Cooten R, Forbes JF. Improving the quality of life during chemotherapy for advanced breast cancer. **N Engl J Med** 1987; 317: 1490–1495.
29. Morris JH, Suissa S, Sherwood S, Wright SM, Greer D. Last days: a study of the quality of life of terminally ill cancer patients. **J Chron Dis** 1986; 39: 47–62.
30. Greer DS, Mor V, Morris JN, Sherwood S, Kidder D, Birnbaum H. An alternative in terminal care: results of the national hospice study. **J Chron Dis** 1986; 39: 9–26.
31. Nakache R, Tydén G, Groth CG. Quality of life in diabetic patients after combined pancreas-kidney or kidney transplantation. **Diabetes** 1989; 38: 40–42.
32. Sage WM, Rosenthal MI, Silverman JF. Is intensive care worth it? An assessment of input and outcome for the critically ill. **Crit Care Med** 1986; 14: 777–782.
33. Blackburn GL, Thornton PA. Nutritional assessment of the hospitalized patients. **Med Clin N Am** 1979; 63: 1103–1115.

34. McGavin CR, Cripta SP, McHardy CJR. Twelve-minute walking test for assessing disability in chronic bronchitis. **Br Med J** 1976; 1: 822–823.
35. Guyatt GH, Pugsley SO, Sullivan MJ *et al.* Effect of encouragement on walking test performance. **Thorax** 1984; 39: 818–822.
36. Lachin JM. Introduction to sample size determination and power analysis for clinical trials. **Contr Clin Trials** 1981; 2: 93–113.
37. Cropp A, DiMarco AF. Effects of intermittent negative pressure ventilation in respiratory muscle function in patients with severe chronic obstructive pulmonary disease. **Am Rev Resp Dis** 1987; 135: 1056–1061.
38. Gutierrez M, Beroiza T, Contreras G, Diaz O, Cruz E, Moreno R *et al.* Weekly cuirass ventilation improves blood gases and inspiratory muscle strength in patients with chronic airflow limitation and hypercarbia. **Am Rev Resp Dis** 1988; 138: 617–623.