Technology Assessment

Technology Assessment Program

Pulmonary Rehabilitation for COPD and other lung diseases

Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, Maryland 20850

November 21, 2006

Pulmonary Rehabilitation for COPD and other lung diseases

Thomas A. Trikalinos, MD

Gowri Raman, MD

Bruce Kupelnick, BS

Priscilla W. Chew, MPH

Joseph Lau, MD

November 21, 2006

Tufts-New England Medical Center EPC

Table of Contents

Table of Contents 3
Summary 5
Background 8
Statement of work 10
Adopted terminology for pulmonary rehabilitation and its components 11
Delineation of key questions and specific tasks 12
Key Questions 12
Key Question 1 12
Key Question 2 13
Methods 16
Definitions 16
Search strategy 17
Inclusion and exclusion criteria 17
Patient population 17
Intervention 18
Comparisons of interest 18
Eligible outcomes 20
Meta-analyses and subgroup or sensitivity analyses 23
Selection of systematic reviews to address key questions 24
Assessing individual randomized controlled trials 25
Applicability and quality of randomized controlled trials 25
Evidence tables and summary tables of reviewed randomized controlled trials 26
Results 27
Key Question 1: Efficacy and safety of exercise-based pulmonary rehabilitation 27
Patients with stable COPD 27
Trials published after the Cochrane systematic review 37
Patients after acute exacerbation of COPD 39
Systematic review 39
Subsequent randomized controlled trials 41
Patients with non-COPD lung disorders 43
Stable patients 43
Patients weaning from mechanical ventilation 44
Subquestion 1.1: Long term effects of pulmonary rehabilitation 47
Subquestion 1.2: Relationship between pulmonary rehabilitation-associated harms and comorbid conditions 50
Subquestion 1.3: Patient level features that modify the effect of pulmonary rehabilitation. 50
Subquestion 1.4: Comparison of pulmonary rehabilitation with general versus individually targeted exercise 50
Subquestion 1.5: Comparison of pulmonary rehabilitation in different settings and of supervised versus unsupervised pulmonary rehabilitation 52
Subquestion 1.6: Efficacy of repeated pulmonary rehabilitation programs 53
Subquestion 1.7: Efficacy and safety of long term maintenance interventions for pulmonary rehabilitation effects 54

Key Question 2: Assessment of specific components in exercise-based pulmonary rehabilitation interventions 58
Subquestion 2.1: Incremental efficacy and safety of exercise training 58
Subquestion 2.2: Efficacy and safety of exercise training compared with other non-exercise PR components 63
Exercise training versus inspiratory muscle training 63
Exercise training versus education 64
Exercise training versus breathing exercises 64
Exercise training versus phone follow-up 64
Subquestion 2.3: Incremental efficacy and safety of non-exercise pulmonary rehabilitation components 68
COPD patients 68
Patients with bronchiectasis 73
Subquestion 2.4: Efficacy and safety of different modes of exercise training 73
Higher versus lower intensity training 74
Endurance versus strength training 74
Continuous versus interval training 83
Overview and conclusions on the efficacy and safety of pulmonary rehabilitation 85
Efficacy of exercise-based pulmonary rehabilitation versus conventional care 86
Safety of exercise-based pulmonary rehabilitation versus conventional care 87
Relative value of different exercise training protocols and of different pulmonary rehabilitation components 88
References 90
Appendices 99
Appendix A. Detailed search strategy 101
Appendix B. Comments on validated outcomes 103
Chronic Respiratory Disease Questionnaire (CRDQ) 103
St George's Respiratory disease Questionnaire (SGRQ) 104
6 minute walk test (6MWT) 104
References to Appendix B. 105
Appendix C. The Jadad quality scale 106
References to Appendix C. 106
Appendix D. Evidence tables for randomized controlled trials not included in the published systematic reviews used in this report 107

Summary

Pulmonary rehabilitation (PR) is a multidisciplinary and comprehensive intervention for patients with chronic respiratory diseases who are symptomatic, and often have decreased daily activities. Candidate patients for PR may have chronic obstructive pulmonary disease (COPD), asthma, bronchiectasis, ventilator dependency or other diseases. Integrated into the individualized patient care, PR is hypothesized to reduce symptoms, optimize functional status, increase participation, and reduce health care costs through stabilizing or reversing systemic manifestations of the disease. However, PR is not hypothesized to reverse deranged pulmonary mechanics.

This technology assessment is based on a systematic review of the scientific literature and focuses on randomized controlled trials (RCT) or meta-analyses thereof. The key questions that it addresses were formulated by the Agency for Healthcare Research and Quality (AHRQ) and the Centers for Medicare and Medicaid Services (CMS). Briefly, the technology assessment summarizes the available evidence on the efficacy and safety of PR interventions, and describes the influence of patient-level or study-level characteristics.

Exercise training is the cornerstone of PR. For operational purposes we defined PR as any intervention that included an exercise-training component of at least 2 weeks' duration and optionally one or more non-exercise components: educational, psychosocial support, breathing exercises, respiratory muscle training, or nutritional interventions. We placed emphasis on trials applicable to the Medicare population.

Overall this technology assessment is based on a re-analysis of 44 RCT included in three published systematic reviews, and 26 additional RCT that had not been assessed by these reviews. There is little evidence on the effects of PR in diseases other than COPD. The two

eligible trials in non COPD population yielded results similar to that obtained from the COPD trials.

Overall, exercise-based PR is effective in improving the patients' disease-specific health-related quality of life, as well as their functional and maximal exercise capacity. Especially in the short term, the improvements are significantly larger than the minimal clinically meaningful improvement. Moreover, evidence suggests that exercise-based PR interventions may reduce hospitalizations and primary care consultations. However, these effects are not translated in survival benefits at least among patients with stable COPD. There is also evidence favoring exercise-based PR among patients recovering from or recently recovered from acute exacerbations of COPD. The assessed RCT did not provide evidence on the safety of PR interventions and on which comorbid conditions predispose patients to/or protect patients from adverse events.

Most of the trials were small and many of them have major methodological shortcomings. Analyses of these trials failed to identify statistically significant differences between PR protocols that included only exercise training versus protocols that also included additional, non-exercise-based, components. The same was true when we compared PR protocols that were tailored to address each patient's specific weaknesses versus PR protocols that were common to all patients; and PR protocols focusing on strength training versus protocols focusing on endurance training or combined strength and endurance training. We should caution that absence of statistically significant findings in the aforementioned comparisons does not imply equivalence of the pertinent PR protocols, and should not be interpreted as such.

Based on few small trials with methodological shortcomings, there is insufficient evidence to draw robust conclusions on whether exercise training has an incremental impact when added to non-exercise PR components like education or inspiratory muscle training, or not.

We did not find statistically significant differences when we compared exercise training alone with non-exercise components alone., we did not find statistically significant differences when we assessed the incremental impact of non-exercise components when added to exercise training. However, we stress that all the aforementioned results should be viewed with caution because they are based on few studies of the limited sample sizes methodological shortcomings.

Background

Pulmonary rehabilitation (PR) has been defined in a 1999 joint statement of the American Thoracic Society (ATS) and the European Respiratory Society (ERS) as a "multi-disciplinary program of care for patients with chronic respiratory impairment that is individually tailored and designed to optimize physical and social performance and autonomy".[1] The recent update of that statement considers PR as an "evidence-based, multidisciplinary, and comprehensive intervention for patients with chronic respiratory diseases who are symptomatic and often have decreased daily life activities. Integrated into the individualized treatment of the patient, pulmonary rehabilitation is designed to reduce symptoms, optimize functional status, increase participation, and reduce health care costs through stabilizing or reversing systemic manifestations of the disease".[2]

PR interventions consist at minimum of some form of exercise training, and most commonly of a variety of additional interventions (education, breathing exercises and respiratory muscle training, nutritional interventions, psychosocial support). The aim of PR is to improve the perceived health-related quality of life (HRQOL) of patients with chronic lung diseases by reducing the severity of their symptoms, their disability and handicap, and by improving their functional independence.[1;3-5]

Patients with chronic lung disease, especially patients with chronic obstructive pulmonary disease (COPD) may experience marked dyspnea and intolerance to exercise.[1;5] It is postulated that these symptoms of lung disease are also related to the patients' generalized weakness and their comorbidities (e.g., cor pulmonale; heart failure, renal failure, etc).[1;5] Dyspnea in turn results in reduced activity, which worsens and perpetuates muscle weakness, and a vicious cycle between muscle de-conditioning and shortness of breath is established. This

is especially true for ambulatory muscles. In addition, many everyday tasks are performed using upper limb and torso (shoulder, neck and abdominal) muscles, some of which act as accessory respiratory muscles. During such tasks the reduced contribution of accessory respiratory muscles to ventilation may intensify shortness of breath. Notwithstanding symptoms directly related to exertion, patients with chronic respiratory disease may experience depression, anxiety and social isolation.

PR has been employed in the attempt to reverse some of these aforementioned pathophysiological and psychosocial conditions. Exercise training aims to improve muscle strength and endurance, and to optimize their use by the patients. It is also hypothesized to decrease general fatigue. Participation in PR programs is expected to reduce anxiety and social isolation. This may be a result of the anti-depressive effects of exercise, and of the beneficial role of education and other interventions apart from exercise training.

The major national and international respiratory organizations (ATS/ERS,[1;2] the American College of Chest Physicians [ACCP] jointly with the American Association of Cardiovascular and Pulmonary Rehabilitation [AACVPR],[3] and Global initiative for chronic Obstructive Lung Disease [GOLD, www.goldcopd.com]) have recommended PR as the standard of care in the treatment of moderate to severe chronic respiratory disease.

A large number of randomized controlled trials (RCT) have been reported on PR interventions in participants relevant to the Medicare population. The majority of these trials evaluated patients with COPD. There is a dearth of evidence for exercise-based PR (randomized trials in particular) on patients with other lung diseases. The role of non-exercise components of PR interventions has not been extensively studied in COPD patients.

Statement of work

The Center for Medicare and Medicaid Services (CMS) has requested a technology assessment through the Agency for Healthcare Research and Quality (AHRQ) on PR primarily for COPD and conditions such as asthma, bronchiectasis, ventilator dependency, and other relevant respiratory illness. The objective is to address specific questions about safety and effectiveness of PR. There seems to be limited evidence on safety and effectiveness of PR for other conditions of interest, apart from COPD. Also, the evidence of PR effectiveness in the elderly has not been systematically evaluated. Specific components of PR and subgroups of patients eligible for PR are also of interest to CMS.

The overarching question of interest to CMS is: What is the evidence for safety and effectiveness of PR for patients ≥65 years old with COPD, asthma, bronchiectasis and other relevant conditions? CMS is also interested in a description of the outcomes measures reported in the studies, a summary of the evidence on complications, harms, and adverse events associated with PR that have been reported, and an assessment on whether conditions prevalent in the older Medicare population increase the risk for these events with PR. Specific factors of interest to CMS include:

a. Internal and external validity of the studies (includes inclusion and exclusion criteria of the studies).
b. Length of follow-up
c. Intensity of treatment, number and frequency of sessions
d. Patient characteristics (i.e., gender, comorbidity) and disease characteristics (i.e., disease severity). Age of patients and generalizability to Medicare population.
e. Concurrent treatment with β-agonists and/or hormonal treatments and/or new therapies (e.g., spiriva) and/or concurrent treatment with supplementary oxygen
f. Concurrent PR in disease management programs
g. Place of delivery (e.g., home, inpatient, outpatient).
h. Physician supervision

i. Components of the PR and whether components were individually tailored or generalized
j. Persistence of benefits/harms over time
k. Repeated course of PR

Adopted terminology for pulmonary rehabilitation and its components

Exercise training is commonly perceived as the cornerstone of PR, and is considered by many a *sine-qua-non* for PR interventions.[1;3;5-10] In this technology assessment we define and adopt the following terms:

- *Exercise training* is any intervention that focuses on endurance or strength training of large skeletal muscles, like upper limb and lower limb muscles. Distinct types of exercise training interventions exist. "Passive" training of large skeletal muscles with electrostimulation is not included in the definition of exercise training.
- *Inspiratory muscle training* (IMT) refers to training of the inspiratory muscles through resistance breathing, threshold breathing or isocapnic hyperpnea (volume training). Based on expert input and CMS input, IMT is the only form of respiratory muscle training that was assessed.
- *Non-exercise PR components* stand for educational, psychosocial or behavioral interventions (alone or in combinations).
- *Conventional care* is a broad term. It refers to the care that different patient subgroups receive in current clinical practice. Conventional care is different for people with stabilized disease compared with patients after an acute exacerbation of chronic disease. For example, patients in respiratory intensive care units may routinely receive some form of exercise training during weaning from ventilators.

- *Usual community care* refers to the conventional care that non-hospitalized patients with chronic respiratory diseases receive in current clinical practice, and implies absence of interventions with exercise training and/or non-exercise components.
- *Exercise-based PR* refers to any PR intervention that *at minimum* has an exercise-training component. Exercise-based PR may include non-exercise PR components or other supplemental interventions on top of exercise training.

Delineation of key questions and specific tasks

We reorganized the statement of work into specific Key Questions (and subquestions). The influence of the "factors" (study-level or patient-level characteristics) mentioned in the statement of work on the efficacy of PR interventions can be assessed directly and indirectly. The indirect assessment compares RCT of PR with a "control" intervention with respect to a specific factor (e.g., among RCT of PR versus "control", contrast those that used high or low intensity of exercise training). The direct assessment is may be more preferable, and pertains to RCT that directly compare different intervention protocols. The direct assessment is applicable only to some of the above factors that refer to variations of PR protocols: One may for example seek RCT that directly compare high and low intensity exercise training. Some of the specific subquestions we have posed emphasize this second option of head-to-head comparison of different PR protocols.

Key Questions

Key Question 1

What is the efficacy and safety of PR for patients in the Medicare population aged ≥65 years who have COPD, asthma, bronchiectasis or other relevant conditions?

All subquestions under the key questions place emphasize specific areas of interest. Subquestions 1.1 to 1.3 can be assessed by RCT that compared exercise-based PR versus conventional care.

KQ1.1. What are the long term effects of PR?

KQ1.2. Is the risk for PR-associated complications, harms and adverse events increased by comorbid conditions prevalent in the older Medicare population?

KQ1.3. Are there patient level features (i.e., characteristics pertaining to individual patients, like gender, diagnosis, cotreatments, and similar) that modify the effect of PR?

The following subquestions (1.4 to 1.7) can be assessed by specific RCT that compare different protocols of exercise-based PR. We pose them to facilitate organizing:

KQ1.4. What is the efficacy and safety of PR with general exercises compared with PR with individually targeted exercise?

KQ1.5. What is the efficacy and safety of PR interventions that are located in different settings (home-based, outpatient, inpatient), and of supervised compared with unsupervised PR?

KQ1.6. What is the efficacy and safety of repeated programs of PR?

KQ1.7. What is the efficacy and safety of long term maintenance interventions for PR?

Key Question 2

What is the efficacy and safety of specific PR components in exercise-based PR interventions? There are many subquestions that could be posed, because there are many different combinations of PR components. With input from CMS we decided to focus on comparisons where exercise training was part of at least one of the comparators. Thus, we assess the value of non-exercise components with respect to exercise training. We use comparison schemes that, in principle, allow the isolation of the effects of different PR components.

Table 1. Classification of the different comparison schemes selected for evaluation.

Comparison	***Comparator A***	***Comparator B***	***This comparison provides answers on:***
1	Exercise training + non-exercise PR component(s)	Same non-exercise PR component(s)	*Exercise training*. Assesses the incremental effect of exercise training
2	Exercise training	Non-exercise PR component	*Head-to-head comparison*
3	Exercise-based PR + non-exercise PR component(s)	Same exercise-based PR	*Non-exercise PR component(s).* Assesses the incremental effect of non-exercise PR component to an exercised-based PR program.
4	Exercise training protocol with characteristic A (see comments column)	Exercise training protocol with characteristic B	*Different exercise training protocols:* Higher or lower intensity of training; endurance or strength training; and continuous or interval training

Based on the above, we devised the following specific subquestions, which can be assessed by specific RCT to facilitate organization:

KQ2.1. What is the incremental efficacy and safety of exercise training when added to non-exercise PR components? (assessed by comparison 1 in Table 1)

KQ2.2. What is the efficacy and safety of exercise training compared with other non-exercise PR components? (assessed by comparison 2 in Table 1)

KQ2.3. What is the incremental efficacy and safety of non-exercise PR components when added to exercise-based PR? (assessed by comparison 3 in Table 1)

KQ2.4. What is the efficacy and safety of different modes of exercise training, specifically:

a. Higher versus lower intensity training?

b. Endurance versus strength exercise training?

c. Continuous versus interval training?

(assessed by comparison 4 in Table 1)

Methods

This technology assessment is based on a systematic review of the literature. We retrieved published systematic reviews and meta-analyses, whenever such were available to identify potential relevant studies. A MEDLINE search was conducted to identify additional RCT that address questions not covered by the systematic reviews, or that update existing eligible systematic reviews.

Definitions

Disease severity was defined for COPD only. It was not defined for other diseases because of the dearth of RCT on other diseases. The BODE index (body-mass index (B), degree of airflow obstruction (O) and dyspnea (D), and exercise capacity (E), measured by the 6 minute walk test) is a composite 10 point score that was proposed in 2004 to predict clinical deterioration and survival in people with COPD.[11] However, the vast majority of the retrieved research was completed before its introduction. We therefore classified COPD severity according to the GOLD classification scheme.[12] Because patient level data were not available, we based our classification on the average FEV_1 values in each RCT. Hence, GOLD I was $FEV_1 \geq 80\%$ of predicted; GOLD II $80\% > FEV_1 \geq 50\%$ of predicted; GOLD III $50\% > FEV_1 \geq 30\%$ of predicted; and GOLD IV <30% of predicted. When absolute FEV_1 values instead of proportions were given, we considered that $FEV_1 < 1.0$L would fall in the GOLD IV category, and that the boundary between GOLD II and III would be 1.5L.

We defined a trial as long term if it followed patients for 12 months or more. We characterized PR programs of more than 12 weeks duration as long duration interventions.

Search strategy

We conducted a comprehensive search in MEDLINE from its inception through April 25, 2006 to identify English language publications of RCT, systematic reviews, and meta-analyses of PR among adults. Search terms included the following pulmonary disease terms: lung diseases,chronic obstructive pulmonary disease, asthma, bronchiectasis, lung transplantation, and artificial respiration; and the relevant PR interventions terms: rehabilitation, exercise therapy, exercise movement techniques, exercise tolerance, and physical therapy modalities; and the relevant study designs (for details refer to the complete search strategy in appendix A).

Inclusion and exclusion criteria

After consultation with the technical expert, AHRQ and CMS, we considered all research publications meeting the criteria described in the following sections. We considered all relevant RCT and systematic reviews and meta-analyses. We did not review RCT that were published only as abstracts, or meta-analyses in which all of the included RCT were published as abstracts. Papers published in languages other than English were also excluded (unless they were included in a systematic review along with trials published in English).

Patient population

Eligible research reports focused on patient groups with mean age of 59 years or older, who had COPD or another chronic respiratory disease of interest such as asthma, lung cancer, bronchiectasis, interstitial lung disease (idiopathic) and chest wall disease. We excluded peri-operative PR interventions (defined as less than 3 months before or after a major operation such as abdominal or chest surgery), post-polio syndrome and muscular dystrophies, patients with

tetraplegia/spinal injury, patients with cystic fibrosis and patients with more rare diseases unlikely to impact on a population level (e.g., lung disease in scleroderma).

Intervention

We considered eligible research where the intervention was an exercise-based PR intervention and consisted of 2 weeks or longer of any exercise training for large skeletal muscles with at least two sessions per week, with or without any non-exercise PR component. The comparator could be any intervention. We excluded all research where exercise–based PR had not been administered in any patient group. We excluded trials in which the intervention was "passive" training of large skeletal muscles with electrical stimulation, or where the experimental intervention was yoga or tai chi exercises.

Comparisons of interest

We considered eligible research that addressed the following comparison schemes:

- Exercise training ± non-exercise PR component(s) compared with conventional care.

  Here conventional care would be usual community care for patients with stable chronic disease, or conventional treatment for patients in the respiratory care unit weaning from ventilator dependency. Whenever we encountered RCT with more than two arms, the arm with the more comprehensive PR intervention (exercise training plus the more non-exercise PR components and/or IMT) was compared with conventional care.
- Exercise training + non-exercise PR component(s) compared with the same non-exercise PR component(s), or

Exercise training compared with non-exercise PR component(s), or

Exercise-based PR + non-exercise PR component(s) compared with the same type of exercise-based PR

From these comparison schemes we excluded RCT where the non-exercise PR components were *supplemental interventions*. These were defined as pharmacological (including O_2 supplementation during exercise, tiotropium administration during the PR etc), nutritional (e.g.,polyunsaturated fatty acids administration) or other interventions (e.g.,ventilation support) aiming to facilitate or enhance the effects of exercise training. CMS deemed that the comparison of PR with supplemental interventions versus PR without supplemental interventions does not need to be assessed, on the basis that supplemental treatments are largely non-covered and cannot easily be covered by Medicare.

- Exercise-based PR with a given exercise training protocol compared with same PR with different exercise training protocol with respect to the following parameters:
    - General exercises compared with PR with individually targeted exercise: Whether all patients received the same exercises versus whether patients received specific exercises that addressed their personal weaknesses or needs.
    - Setting of the PR intervention (home; community; hospital).
    - Whether intervention was supervised by a health care professional versus the same intervention unsupervised.
    - The intensity of exercise training: Whether patients exercised at levels ≥60% of their personal maximum workload (higher intensity) versus

lower intensity exercise, while keeping the other exercise training characteristics (frequency and duration) reasonably similar.

- Endurance versus strength training: Exercise training protocol included only endurance training (aerobic exercises) versus protocols that included only strength training, or added strength training to endurance training.
- Continuous exercise versus interval exercise training: Whether exercises were performed in a continuous manner at a given, constant intensity versus exercise training consisting of brief bouts of high (or maximal) intensity exercise separated by periods of lower intensity exercise (or rest).

Especially for the last three factors (i.e., high or low intensity, endurance or strength and continuous or interval training) we required that all exercise training protocols be standardized, as previously suggested.[9] Standardized exercise was defined as identical exercise training schemes (treadmill walking, cycle ergometer or weightlifting training) at intensities that are objectively measurable (e.g., in Watts). For example outdoor walking "at a quiet pace" would not qualify as standardized exercise training.

- After the end of an exercise based PR intervention, presence compared with absence of repeated PR or other maintenance strategy.

Eligible outcomes

Outcomes were assessed at longest follow-up, unless otherwise stated.

Effectiveness and efficacy

Where applicable, we focused on the description of outcome measures in their natural units (e.g., meters, Watts, minutes), rather than relative percentage changes from an average reference value. Main metric of interest was the difference in the changes from baseline for continuous measures and the odds ratio or the risk ratio for binary outcomes. We selected the following outcomes as important in the evaluation of the effectiveness or efficacy of exercise-based PR:

- The main dyspnea-specific outcome was the baseline dyspnea index/transitional dyspnea index (BDI/TDI) score.[13] Other were the Visual Analog Scale (VAS), the modified Borg scale, and the Chronic Respiratory Disease Questionnaire (CRDQ) dyspnea domain score.[14]
- Disease specific Quality of Life (QoL) outcomes: the four domains of CRDQ (dyspnea, mastery, fatigue, emotion)[14] and the total score in St George's Respiratory disease Questionnaire (SGRQ).[15]
- Exercise capacity was measured as
    - maximal exercise capacity (reported outcomes on incremental exercise tests like incremental cycle ergometry test, ICET, incremental treadmill walking and incremental shuttle walking test [ISWT] results) with main outcome the maximal workload, W_{max}
    - functional/endurance exercise capacity (e.g.,6 minute walk test [6MWT] results, as distance in meters walked in 6 minutes).
    - endurance at constant work rate (e.g., endurance time at 50% of maximal exercise capacity in a treadmill walking test)

Note that the assessment of maximum exercise capacity with incremental protocols on a cycle ergometer or a treadmill, as well as the assessment of endurance capacity with constant work rate tests on a cycle ergometer or a treadmill are "laboratory-based" tests. In contrast the 6MWD and the ISWT that assess the functional capacity are mainly "field-based" tests, conducted outside the laboratory.

- General outcomes include all cause mortality; and (re)admission rates.
- Inspiratory capacity

Safety

Adverse events, as defined in the primary reports, and all cause mortality were considered as a safety outcome. All safety outcomes were assessed at longest available follow-up.

Dropouts are clearly related to the efficacy of the interventions. Patients who perceive the treatment as non-efficacious may opt to leave the study early. Dropouts are also influenced by many practical and logistic issues (e.g., personal time required for study participation, transportation issues, loss of interest), and therefore were not used as a proxy for adverse events or as a safety outcome.

Outcomes not analyzed

PR cannot reverse the derangement of pulmonary mechanics (as assessed by FEV_1, other lung volumes and lung capacities) and this is especially true for COPD.[1;3;4] Thus, changes in pulmonary physiologic measurements were not reviewed; a possible exception was the lungs' inspiratory capacity (IC), which was suggested by the technical expert as a potentially useful index with direct clinical meaning for people with COPD.

We decided not to review changes in other, additional physiological measurements, despite the fact that they may convey valuable ergophysiological or other specialized information. Thus, after discussions with the technical expert, we concluded that we would not review quantities like the maximum inspiratory pressure (PI_{max}), maximum expiratory pressure (PE_{max}), endurance time of respiratory muscles, heart rate, respiratory frequency, peak O_2 uptake during exercise, muscle strength, muscle thickness or other similar quantities.

Meta-analyses and subgroup or sensitivity analyses

We performed meta-analyses whenever two or more trials where available on the same question, regardless of the extent or significance of between-study heterogeneity. We combined continuous measurements using inverse variance random effects models.[16] The summary continuous effect sizes were expressed as weighted mean differences (WMD) expressed in their natural units. For binary outcomes we combined odds ratios or risk ratios using the DerSimonian and Laird random effects method.[17] Random effects models allow for between-study heterogeneity (dissimilarity) in the meta-analyses and incorporate it in the calculations. We tested for heterogeneity in meta-analyses with Cochran's *Q* statistic (considered significant at $p<0.10$ [16]) and quantified its extent with the I^2 statistic.[18] I^2 ranges between 0 and 100% and higher values imply greater heterogeneity.

We performed subgroup analyses with respect to several factors mentioned below. These were implemented in a random effects meta-regression framework,[19] provided that more than 5 trials were available, so as to avoid overfitting. We hypothesized that trials with higher internal validity may be associated with less impressive findings in the outcomes of interest (RCT with blinded versus non-blinded assessors; RCT with quality grading A versus B or C [see section

below on quality grading]). We also assessed in meta-regressions the following factors, which we considered possibly associated with better outcomes: interventions with duration longer than 12 weeks versus interventions of 12 weeks or less; interventions including non-exercise PR components versus exercise training only; supervised versus unsupervised PR; home-based versus in- or out-patient exercise training; and patients with moderate disease (GOLD II and III) versus patients with severe disease (GOLD IV). We considered that there was evidence for subgroup effects only when the meta-regressions yielded statistically significant inferences.

Where applicable, we performed sensitivity analyses by excluding RCT published in languages other than English and assessing whether inferences changed. For some papers where numerical information was given in figures, rather than text or tables, we extracted the needed information by electronically digitizing the corresponding plots with specialized software (Engauge digitizer ver 2.12, Mark Mitchell). Meta-analyses were performed in Intercooled Stata 8.2 (Stata Corp. College Station, TX)

Selection of systematic reviews to address key questions

In discussions with AHRQ and CMS we decided to capitalize on existing systematic reviews and meta-analyses, provided that they did not have serious methodological flaws, and that they addressed at least one of the comparisons we have delineated in the inclusion/exclusion criteria in an eligible patient subgroup (see below on the assessment of existing systematic reviews). Using these systematic reviews/meta-analyses as a starting basis, we reviewed any additional relevant RCT published after these systematic reviews. We retrieved the individual RCT described in the systematic reviews and perused them to clarify important details when needed.

The identified systematic reviews and meta-analyses of RCT were critically evaluated. We required that eligible systematic reviews assessed comparisons similar to the ones delineated in our inclusion and exclusion criteria. High quality systematic reviews had a clearly described and sufficiently thorough search strategy (ideally searching more than one electronic databases and using an alternative searching algorithm such as hand searching of selected journals); had clearly described inclusion and exclusion criteria; had assessed the internal and external validity of the reviewed RCT according to standard, published methods; clearly described the methods used for meta-analysis and appropriately employed them; and reached conclusions supported by their data. If more than one such systematic reviews or meta-analyses were identified, we retained the one that was reported in greater detail and addressed specific questions that were closer to our key questions.

Assessing individual randomized controlled trials

Applicability and quality of randomized controlled trials

The applicability of RCT was assessed based on disease spectrum, the setting of the interventions (outpatient or home-based rather than inpatient for non-critically ill patients), the gender distribution of the participants, and the selection criteria of the individual RCT (presence of a run-in period; focus on patient subgroups). The inclusion criteria ensured that the average age of the participants was similar to that of the Medicare population.

We assessed the methodological quality of RCT based on whether or not they clearly reported specific quality items: description of randomization, blinding of test assessors (blinding of patients is not feasible), concealment of patient allocation, presence of ad-hoc power analyses, and description and magnitude of attrition rates. We classified the methodological quality of

RCT in to a three-point scale (A, B, C or good, moderate, poor, respectively, as described below). For RCT described in systematic reviews, we relied on the reviewers' assessments, and translated them in our three-point scale.

Grade A (good methodological quality) studies fulfill most commonly held concepts of high quality, including the following: a formal randomized study; clear description of the population, setting, intervention and comparison groups; appropriate measurement of outcomes; appropriate statistical and analytic methods and reporting; no reporting errors; not excessive dropouts (<20%); clear reporting of dropouts; and no obvious bias.

Grade B (moderate methodological quality) studies may be susceptible to some bias, but not sufficient to invalidate the results. Such studies do not meet the criteria described in category A. They have some deficiencies but none likely to cause major bias. Study may be missing information making assessment of the limitations and potential problems difficult.

Grade C (poor methodological quality) studies are subject to significant bias that may invalidate the results. Such studies may have serious errors in design, analysis or reporting. These studies may have large amounts of missing information or discrepancies in reporting.

Evidence tables and summary tables of reviewed randomized controlled trials

Evidence tables for all eligible RCT that were individually reviewed are provided in Appendix D.

We constructed summary tables that capture the important information from the included meta-analyses and RCT.

Results

The literature search identified 2200 citations. After initial screening, 158 citations were considered potentially relevant and full text articles were retrieved. Screening reference lists of reviews and meta-analyses yielded 8 additional papers. Overall, we included 3 systematic reviews (that collectively included 44 RCT), and 26 additional RCT that had not been assessed in the 3 systematic reviews. The majority of the RCT that were already included in the systematic reviews have to be re-reviewed to variable extent to clarify important details when needed.

Key Question 1: Efficacy and safety of exercise-based pulmonary rehabilitation

Patients with stable COPD

We identified six systematic reviews/meta-analyses addressing the efficacy and/or safety of PR in patients with stable COPD.[6;7;20-23] Of these, five articles were rejected for the following reasons: One was completely covered and updated by the Cochrane review.[6] Another, devoted very limited space to the assessment of PR interventions because it aimed at a general overview of contemporary COPD management, and was therefore deemed to be less informative.[23] The third[21] considered randomized and quasi-randomized trials together, and had employed a different definition of eligible exercise-based PR. The fourth meta-analysis[20] focused on the presentation of effect sizes in a unit-free scale, assessed RCT quality using questionable quality items and had different definition of eligible exercise-based PR compared to ours. Finally, the fifth focused only on people with mild-moderate COPD (and thus excluded several eligible studies).[22]

We included one Cochrane review published in 2002 (last non-substantial update in 2004) that assessed the short term effectiveness of exercise-based PR interventions compared to usual community care.[7] The authors scrutinized 23 RCT on patients with stable COPD published between 1977 and 2000. We re-analyzed the 19 (out of 23) trials that provided quantitative data and were eligible according to our inclusion criteria in the meta-analyses described below (Tables 2 through 4).

Overall, the included trials were small, as they randomized a median of 50 participants in the compared arms. Reporting of methodological quality items in most trials was moderate to poor. The Cochrane review used the Jadad scale to score the methodological quality of the individual RCT (see appendix C for details on the Jadad scale). The maximum score in the Jadad scale is 5. However, trials on PR interventions cannot be double blinded, and thus the maximum possible Jadad score for PR trials is 4 (double blinding is an item in this quality scale – Appendix C). No trials scored 4 (out of 4) and only six trials scored 3 (out of 4) on the Jadad scale. Exercise training of ambulatory muscles was included in all analyzed trials, and seven trials included training of the upper limbs. The trials varied in the duration, frequency and intensity of training. Overall, endurance training was evaluated in 18 RCT, with additional strength training in four trials.

The included RCT were deemed to have good applicability. The majority of the participants were males (median 71%). The average age ranged from 60 to 73 years, and baseline average FEV_1 values ranged from 26% to 60% of predicted. Five trials enrolled participants with severe COPD and 13 trials with moderate and/or severe COPD.

Dyspnea and disease-specific QoL (CRDQ scores)

Nine trials (277 patients in experimental arm and 242 in the usual community care arm) used the CRDQ instrument.[24-31] All assessed the effects on the CRDQ-dyspnea score, and eight assessed all four CRDQ domains (Table 2). There was a clinically meaningful improvement favoring the PR arm in the CRDQ dyspnea, fatigue and mastery of breath domains (i.e., difference at of least 0.5 units in a 7 point scale). The corresponding weighted mean differences in the average change from baseline were 1.0 (95% confidence interval: 0.8, 1.2), 0.9 (95% confidence interval: 0.7, 1.1) and 0.9 (95% confidence interval: 0.7-1.2). The weighted mean difference of the emotional function domain was 0.7 (95% confidence interval: 0.4, 1.0), and the confidence interval could not exclude the minimal clinically significant difference of 0.5 units. There was no statistically significant heterogeneity among the RCT included in the meta-analyses, nor statistically significant differences across subgroups.

Table 2. Randomized controlled trials of pulmonary rehabilitation versus usual community included in the Cochrane review: Chronic respiratory disease questionnaire (short term)

Study, year	**N_r**	**COPD severity**	**Male %**	**Intervention**		**CRDQ (95% CI), on a seven point scale**					*Quality*
				Exercise; frequency; setting	**Other components**	**N_a (%)**	**Dyspnea**	*Fatigue*	*Mastery of breath*	*Emotional function*	
Cambach, 1997[a,b]	23	II	57	LLE (End) + ULE, 3/wk for 12wk; community-based	IMT; Edu	22 (96)	1.2 (0.4, 2.0)	1.3 (0.4, 2.1)	1.3 (0.3, 2.2)	0.4 (-0.5, 1.3)	B
Griffiths, 2000[a]	200	III	54	LLE (End) + ULE, 3/wk for 6 wk; out-patient + home-based	Edu; Psy; Nutr; SmC	184 (92)	1.2 (0.9, 1.5)	1.1 (0.7, 1.5)	1.1 (0.7, 1.4)	1.2 (0.8, 1.5)	A
Goldstein, 1994	89	III	48	LLE (End) + ULE (End), ≥3/wk for 8 wk; in-patient	BE; Edu; Psy	79 (89)	0.7 (0.1, 1.2)	0.4 (-0.2, 0.9)	0.8 (0.2, 1.3)	0.4 (-0.1, 1.0)	B
Troosters, 2000[c]	100	III	87	LLE (End + Str), 3/wk for 12 wk and then 2/wk for 12 wk; out-patient	No	62 (62)	0.8 (0.2, 1.5)	0.7 (0.1, 1.4)	0.9 (0.2, 1.7)	0.6 (-0.0, 1.3)	C
Guell,1995	60	III	100	LLE (End), 5/wk for 12 wk; out-patient	BE; PD	56 (93)	1.3 (0.6, 2.0)	1.1 (0.5, 1.7)	1.2 (0.5, 1.9)	1.0 (0.3, 1.7)	B
Wijkstra, 1994	45	III	86	LLE + ULE, 7/wk for 12 wk; home-based	IMT; BE; Edu; Psy	43 (96)	0.9 (0.1, 1.7)	0.6 (-0.1, 1.4)	0.6 (-0.1, 1.3)	0.5 (-0.1, 1.1)	B
Hernandez, 2000	60	III	100	LLE (End), 6/wk for 12wk; home-based	No	37 (61)	0.8 (0.02, 1.5)	0.9 (0.1, 1.7)	0.7 (-0.3, 1.6)	0.5 (-0.3, 1.4)	C
Simpson, 1992	34	III	54	LLE + ULE (Str), 3/wk for 8wk; out-patient	No	22 (65)	1.2 (0.4, 2.0)	0.8 (-0.1, 1.6)	0.7 (-0.4, 1.8)	0.3 (-0.5, 1.1)	C
Busch, 1988[d]	20	IV	79	LLE (End) + ULE (Str+End), 5d/wk for 18 wk; home-based	No	14 (70)	-0.4 (-2.1, 1.3)	ND	ND	ND	B
Overall, random effects meta-analyses						519	1.0 (0.8, 1.2)	0.9 (0.7, 1.1)	0.9 (0.7, 1.2)	0.7 (0.4, 1.0)	

Participants in all RCT had mean age above 60 years. Quality scoring is based on the Jadad score and the description of allocation concealment.
RCT are sorted by COPD severity (in what would be the GOLD classification equivalent, judged by the average FEV_1 values in cases and controls), and then by size. Heterogeneity was statistically non-significant for all four metrics (p-value [I^2, %] for heterogeneity was 0.53 [0%], 0.48 [0%], 0.87 [0%], and 0.17 [33%] for the four CRDQ scores in the order they are mentioned in the table).
BE: breathing exercises; d: day; CRDQ: chronic respiratory disease questionnaire; Edu: education; End: Endurance training; IMT: inspiratory muscle training; LLE: lower limb exercise; N_a number analyzed and % of randomized; ND: No data; N_r number randomized; PD: postural drainage; Psy: Psychosocial intervention; RCT: randomized controlled trial; SmC: smoking cessation; Str: strength training; ULE: upper limb exercise; wk: week(s)
[a] Male proportion refers to patients after dropouts in all trials except for Cambach and Griffiths where it refers to people randomized.
[b] Refers only to COPD patients; asthmatics in this trial were younger than 59 years on average and are excluded from the technology assessment.
[c] The Cochrane review names this RCT "Gosselink 2000"
[d] Only the CRDQ-dyspnea score is reported in this trial

Maximal exercise capacity

The incremental cycle ergometer test was used to assess the summary effects of PR on the maximal exercise capacity based on 14 trials (255 treated and 233 controls, Table 3).[24;26;28-40] We excluded from the summary estimate two trials, one that measured total work in Joules instead of power (Watts)[24] and one very small French trial that reported a highly unlikely precise estimate.[40] Among the 12 remaining RCT a statistically significant increase in the maximum achieved workload was observed in favor of the PR arm: 7 W (95% confidence interval: 3, 12). It is difficult to interpret the clinical importance of this finding because the minimal clinically meaningful improvement in this test has not been described. There was no statistically significant heterogeneity for this summary estimate and no statistically significant differences across subgroups. The Cochrane review had included the aforementioned trials that were excluded from our analyses; even so the inferences were very similar (weighted mean difference = 5 W [95% confidence interval: 1, 10]).

Table 3. Randomized controlled trials of pulmonary rehabilitation versus usual community included in the Cochrane review: Maximal exercise capacity outcome (short term)

Study, year	N_r	COPD severity	Male %	Intervention		ICET		Quality
				Exercise; frequency; setting	Other components	N_a (%)	in Watts (95% confidence interval)	
Trials we included in the analyses								
Goldstein, 1994	89	III	48	LLE (End) + ULE (End), ≥3/wk for 8 wk; inpatient	BE; Edu; Psy	57 (64)	0.0 (-8.8, 8.8)	B
Troosters, 2000[b]	100	III	87	LLE (End + Str), 3/wk for 12 wk and then 2/wk for 12 wk; out-patient	No	62 (62)	11.0 (-7.3, 29.3)	C
Guell,1995	60	III	100	LLE (End), 5/wk for 12 wk; out-patient	BE; PD	56 (93)	6.4 (-14.2, 26.9)	B
Emery, 1998	55	III	54	LLE (End) + ULE, daily for 5wk; out-patient	Edu; Psy	50 (91)	11.4 (-6.0, 28.8)	B
Engstrom, 1999	55	III	52	LLE (End) + ULE (Str), daily for 52 wk; out-patient	IMT; Edu	50 (91)	8.6 (-5.1, 22.3)	B
Wijkstra, 1994	45	III	86	LLE + ULE, 7/wk for 12 wk; home-based	IMT; BE; Edu; Psy	43 (96)	16.0 (-2.2, 34.2)	B
Hernandez, 2000	60	III	100	LLE (End), 6/wk for 12wk; home-based	No	37 (62)	-5.7 (-23.4, 12.0)	C
Strijbos, 1996	35	III	87	LLE (End), 2/wk for 12 wk; out-patient	PD; Edu; Psy	30 (86)	12.7 (-0.9, 26.3)	B
Simpson, 1992	34	III	54	LLE + ULE (Str), 3/wk for 8wk; out-patient	No	27 (79)	26.3 (-93.4, 146.0)	C
McGavin, 1977	28	III	100	LLE (End), ≥1/d and ≥5/wk, continuous; home-based	No	24 (86)	17.0 (-0.5, 34.5)	B
Jones, 1985	19	IV	50	LLE (End) + ULE, for 10 wk; home-based	No	14 (74)	27.0 (-172.1, 226.1)	B
Lake, 1990	14	IV	71	LLE (End) + ULE, 3/wk for 8 wk; out-patient	No	14 (100)	8.9 (-5.0, 23.0)	B
Overall random effects meta-analysis						464	7.1 (2.5, 11.8)	
Trials we excluded from the analysis (see text)								
Vallet, 1994[a]	20	II	75	LLE (End), for 8 wk; in-patient	BE	12 (60)	0.2 (-0.1, 0.5)	C
Busch, 1988	20	IV	79	LLE (End) + ULE (Str+End), 5d/wk for 18 wk; home-based	No	12 (60)	[different natural quantity][c]	C

Layout as in Table 2. Among the first 12 trials, heterogeneity was statistically non-significant (p=0.59, I^2=0%).
ICET: incremental cycle ergometry test; W: watt(s)
The Cochrane meta-analysis included the last two trials in the table; the results were very similar (5 W [95% confidence interval: 1, 10], p for heterogeneity p-value=0.14 and I^2=30%). Note that the study by Simpson 1992 did not include endurance training.
[a] Published in French.
[b] The Cochrane review names this RCT "Gosselink 2000"
[c] This trial assessed work, not power: 2643.4 J (-1200.0, 6470.7). It was analyzed together with all the other trials in the Cochrane meta-analysis.

Functional exercise capacity (6MWT)

The Cochrane review included ten RCT (235 treated patients and 219 controls) included in meta-analyses for the 6MWT (Table 4).[25;26;29-31;33;34;36;41;42] We included an additional relevant RCT by Bendstrup[43] in our re-analysis, which was omitted from the Cochrane review for non-obvious reasons. The WMD in the 6MWT was 52 meters (95% confidence interval: 37, 67), with marginally non-significant heterogeneity. Excluding Bendstrup et al. makes heterogeneity statistically significant (p=0.08) but its extent is more or less the same (I^2=42%) and the summary results are essentially unchanged. The improvement is not greater than the minimal clinically significant difference in the 6MWT. Hence the clinical significance of this finding is unclear. There were no statistically significant differences across subgroups.

Table 4. Randomized controlled trials of pulmonary rehabilitation versus usual community included in the Cochrane review: Functional exercise capacity outcome (short term)

Study, year	N_r	COPD severity	Male %	Intervention		6 minute walk test		Quality
				Exercise; frequency; setting	Other	N_a (%)	in meters (95% confidence interval)	
Cambach, 1997[a]	23	II	57	LLE (End) + ULE, 3/wk for 12wk; community-based	IMT; Edu	19 (83)	5.0 (-72.2, 82.2)	B
Goldstein, 1994	89	III	48	LLE (End) + ULE (End), ≥3/wk for 8 wk; inpatient	BE; Edu; Psy	77 (87)	43.0 (-2.0, 88.0)	B
Troosters, 2000[b]	100	III	87	LLE (End + Str), 3/wk for 12 wk and then 2/wk for 12 wk; out-patient	No	62 (62)	55.0 (-2.0, 112.0)	C
Guell,1995	60	III	100	LLE (End), 5/wk for 12 wk; out-patient	BE; PD	56 (93)	83.0 (47.9, 118.1)	B
Engstrom, 1999	55	III	52	LLE (End) + ULE (Str), daily for 52 wk; out-patient	IMT; Edu	50 (91)	40.0 (-13.5, 93.5)	B
Wijkstra, 1994	45	III	86	LLE + ULE, 7/wk for 12 wk; home-based	IMT; BE; Edu; Psy	43 (96)	37.0 (-41.3, 115.3)	B
Ringbaek, 2000	55	III	18	LLE (End + Str) + ULE (Str), 2/wk for 8/wk; out-patient	No	36 (65)	29.0 (-24.4, 82.4)	B
Simpson, 1992	34	III	54	LLE + ULE (Str), 3/wk for 8wk; out-patient	No	28 (82)	29.0 (-53.5, 111.5)	C
Booker, 1984	[c]94	IV	ND	LLE (End), 1-2/day for 9wk; home-based	BE; PD; Edu; Psy	69 (73[c])	16.0 (-25.3, 57.3)	B
Bendstrup, 1997[d]	47	IV	56	LLE (End + Str) + ULE (Str), 3/wk for 12wk; out-patient	Edu; OT; SmC	32 (68)	58.2 (22.2, 94.2)	B
Lake, 1990	14	IV	71	LLE (End) + ULE, 3/wk for 8 wk; out-patient	No	14 (100)	143.6 (74.3, 212.9)	B
Overall, random effects meta-analysis						486	50.4 (30.6, 70.2)	

Layout as in Table 2. Heterogeneity was statistically non-significant, but only marginally so (p-value=0.11 and I^2=37%).
ND: Not described; OT: occupational therapy
[a] Refers only to COPD patients; asthmatics in this trial were younger than 59 y on average and are excluded from the technology assessment.
[b] The Cochrane review names this RCT "Gosselink 2000"
[c] This is a three-arm trial. Unclear how many were actually randomized in the two arms, and 94 is an estimate for the 2 arms which have been used here. The percentage of analyzed people (73%) refers to all three arms of this trial.
[d] Bendstrup 1997 is not included in the meta-analyses in the Cochrane review for unclear reasons (Bendstrup 1997 was among the eligible RCT in the Cochrane review). Inclusion or exclusion of this RCT makes no difference in the final estimates. However, without Bendstup 1997 heterogeneity becomes statistically significant (0.08).

Safety

None of the RCT included in the Cochrane review addressed directly any complications secondary to the PR intervention.

Mortality and admission rates

Data on mortality were available on seven RCT (Table 5), after follow-up periods ranging from 3 months to 2 years.[27;30;33;34;38;43;44] There was no overall effect of exercise-based PR on mortality (odds ratio 1.03, 95% confidence interval: 0.54, 1.89), and there was no statistically significant heterogeneity. In one trial the survival status at 18 months was unclear for some of the dropouts.[30] Inferences based on the summary estimate remained similar when all these dropouts were counted as deaths.

Data on acute exacerbations of COPD were systematically recorded in three RCT.[27;34;43] Of these, two trials reported statistically significant fewer exacerbations or hospitalizations (number of admissions and length of stay) in the PR arm compared with the control arm.[27;34] The third trial reported no significant differences across the compared arms.[43] Griffiths[27] reported that participants in the intervention arm had on average more frequent consultations compared to those in the control arm.

Table 5. Randomized controlled trials of pulmonary rehabilitation versus usual community included in the Cochrane review: Mortality

Study, year	N_r	COPD severity	Male %	Intervention		Deaths/Total		Odds ratio (95% confidence interval)	Quality
				Exercise; frequency; setting; follow-up	Other	Intervention	Usual community care		
Griffiths, 2000[a]	200	III	54	LLE (End) + ULE, 3/wk for 6 wk; out-patient + home-based; 12 months	Edu; Psy; Nutr; SmC	6/99	12/101	0.48 (0.17, 1.33)	A
Troosters, 2000[a]	100	III	87	LLE (End + Str), 3/wk for 12 wk and then 2/wk for 12 wk; out-patient; 18 months	No	9/37	7/33	1.19 (0.39, 3.67)	C
Guell,1995	60	III	100	LLE (End), 5/wk for 12 wk; out-patient; 24 months	BE; PD	5/30	3/30	1.80 (0.39, 8.32)	B
Engstrom, 1999	55	III	52	LLE (End) + ULE (Str), daily for 52 wk; out-patient; 12 months	IMT; Edu	2/26	1/24	1.92 (0.16, 22.61)	B
Wijkstra, 1994	45	III	86	LLE + ULE, 7/wk for 12 wk; home-based; 18 months	IMT; BE; Edu; Psy	2/28	1/15	1.08 (0.09, 12.95)	B
Strijbos, 1996	35	III	87	LLE (End), 2/wk for 12 wk; out-patient; 18 months	PD; Edu; Psy	2/18	0/15	4.70 (0.21, 105.79)	B
Bendstrup, 1997[b]	47	IV	56	LLE (End + Str) + ULE (Str), 3/wk for 12wk; out-patient; 3 months	Edu; OT; SmC	1/20	0/22	3.46 (0.13, 89.95)	B
Overall, random effects meta-analysis								1.03 (0.54, 1.89)	

Layout as in Table 2.

Deaths have been calculated with respect to the number of patients followed-up in the corresponding arms, at the longest available follow-up. No statistically significant heterogeneity ($p=0.62$, $I^2=0\%$).

[a] There were 50 randomized patients per arm. 13 and 17 in the intervention and control arm did not attend training sessions or follow-up, and their survival status is unknown. Even if we consider all dropouts dead, the summary odds ratio would be 0.91 (95% confidence interval: 0.54, 1.56), without any heterogeneity ($p=0.62$, $I^2=0\%$)

[b] The number of randomized patients per arm is unclear, because data are reported after 5 post-randomization exclusions.

Trials published after the Cochrane systematic review

We identified three additional RCT from four publications that assessed the efficacy and safety of PR[45-48] (Table 6), all published after the completion of the Cochrane review. The trials enrolled 40 to 191 stable COPD to exercise based PR. One trial was multicenter and assessed PR in the context of a disease management program.[45;46] Two trials blinded outcome assessors.[45-47] The methodologic quality was good in one trial and poor in the other two.[47;48]

Two trials reported significantly greater improvements in favor of the intervention arm in St Georges' Respiratory Questionnaire (SGRQ), but failed to exclude the minimal clinically significant differences in this scale.[45-47] The third trial by Singh reported scores of CRDQ and the results agreed with the Cochrane review.

Measurements in the 6MWT across the compared arms did not differ beyond chance neither at 4 nor at 12 months of follow-up in one trial.[45;46] The trial by Finnerty[47] reported that the improvement in the intervention arm was greater by 51 meters at 12 weeks and 53 meters in the 24 weeks (only the former was significant). The third trial by Singh reported a nonsignificant difference of 48 meters favoring the PR arm.[48]

Reductions in hospitalizations were assessed in Bourbeau only.[45;46] They were significantly reduced in the intervention arm both during the first year (-0.7 per patient per year in the PR arm compared with control) and the second year (-0.44 per patient per year compared with control) of follow-up. Similarly, emergency room visits as well as the length of hospitalization were significantly less for the PR arm. No trials evaluated safety outcomes.

Table 6: Randomized controlled trials published after the Cochrane review (patients with COPD)

Study, year	*Size and applicability*			**PR description**		**Difference in change from baseline effect (95% confidence interval)**		**Quality**
	N_r	**COPD severity**	**Male %**	**Exercise; setting**	**Other component**	**CRDQ (Units)**	**6MWT (meters)**	
Finerty, 2001	55	II	67	LLE (End) + ULE, 5/wk for 6 wk; home-based, out-patient	BE; Edu; Nutr; OT	NA	12 wks: 51m (20 to 81) 24 wks: 53m (p>0.05)	C[a]
Bourbeau, 2003	191	IV	55	LLE (End), 3/wk, continuous; home-based	Edu	NA	No significant changes neither within nor between groups at 4 or 12 months	A
Singh, 2003	40	IV	80	LLE (End), for 4 wks; home-based	BE; PD	At end of treatment • Dyspnea: 0.88 (0.26, 1.50) • Function: 0.75 (0.04, 1.46) • Fatigue: 0.84 (-0.04, 1.72) • Mastery: 0.84 (-0.11, 1.79)	4 wks: 48m (-9, 104)	C[b]

Trials sorted by COPD severity and then by size. As stated in the methods we focused on the same outcomes assessed by the Cochrane review. See text for description of other outcomes.
6MWT: 6 minute walk test; BE: breathing exercises; CRDQ: chronic respiratory disease questionnaire; Edu: education; ICET: incremental cycle ergometry test; ITT: intention to treat; LLE: lower limb exercise; N: number randomized; Nutr: nutritional intervention; OT: occupational therapy; PD: postural drainage; PR: pulmonary rehabilitation; RCT: randomized controlled trial; ULE: upper limb exercise; wk: week(s); y: year(s).
[a] High attrition; suboptimal RCT reporting of methodological quality items
[b] Poor RCT reporting of methodological quality items

Patients after acute exacerbation of COPD

Systematic review

We identified one systematic review on COPD patients that included 6 RCT (140 were randomized to PR and 90 to control) comparing PR after acute exacerbation with conventional care.[10] We excluded an RCT because the duration of exercise training was less than 2 weeks.[49] The remaining five eligible RCT (described in 6 papers) had sample sizes between 26 and 70 patients[50-55] (Table 7). Participants were enrolled after having been hospitalized for acute COPD exacerbations (in one trial both hospitalized and people treated at home were included[53]). Participants' average age ranged from 64 to 70 years and their mean FEV_1 from 32% to 40% of predicted (GOLD categories III to IV). Three fourths of the 70 participants in one RCT[54] needed mechanical ventilation. Across all RCT, the majority (65 to 90%) of participants were male. The maximum duration of PR intervention was 6 months.[50;51;55] Exercise training consisted of aerobic (endurance) and strength exercises in three[52;53;55] and endurance exercises only in the remaining two RCT. Follow-up duration ranged from 5 weeks to 18 months. Reporting of methodological quality items in all six trials ranged from poor to moderate.

Effects on dyspnea

TDI was assessed in two RCT, which reported greater improvements in the PR arm (Table 7). The confidence intervals for the difference in the changes in TDI from baseline excluded the minimal clinically significant difference of 1 point.[56] One other trial evaluated participants who were severely ill and admitted to a respiratory intensive care unit (Nava[54]). It reported a net change of 17mm on a dyspnea visual analog scale (Table 7) favoring of PR. The clinical significance of the PR effects on dyspnea visual analog scale is not known. Two other

heterogeneous trials reported greater improvement in CRDQ-dyspnea in the PR arm compared with the control arm (1.67 [95% confidence interval: 0.36, 1.98].[50-52]

Effects on functional exercise capacity

Three trials[50;54;55] assessed patients after acute exacerbation of COPD with the 6MWT. The summary between-arm difference was 114 m favoring the pulmonary rehabilitation intervention (95% confidence interval:28, 199), but this estimate was very heterogeneous (Table 7). Another two trials[52;53] assessed the differences in the change from baseline in the Shuttle walk test, and found a summary difference of 81m (95% confidence interval: 48, 115) favoring the intervention (Table 7).

Effects on disease specific quality of life

Two trials evaluated disease specific QoL using the CRDQ instrument.[50-52] Overall, their meta-analysis suggests there is a net improvement in all four CRDQ domains with PR. The confidence intervals for the WMD exclude (exceed) the minimal clinically significant difference in this instrument (Table 7), implying a clinically meaningful change in all domains, except for the dyspnea domain. One trial reported that the difference in the mean changes of the total SGRQ score was both clinically and statistically significant[53] in favor of the PR arm.

All cause mortality and hospitalizations

PR appeared to have a protective effect on mortality, with a summary relative risk of 0.40 (95% confidence interval: 0.18, 0.86) with no statistically significant heterogeneity among two trials[52;53] (Table 7). This estimate became non-significant when the trial reported by Nava was added in the calculations.[54] The latter was conducted among severely ill, high-risk patients with

20% mortality in each arm. The follow-up periods of the synthesized trials ranged from 6 weeks to 18 months.

The relative risk for unplanned hospital admissions was 0.26 (95% confidence interval: 0.12, 0.54) in favor of PR in 3 trials[50-53] with no statistically significant heterogeneity.

Safety

Two trials (Man and Behnke[50-52]) specifically stated that any adverse events would be recorded. No adverse events of PR were observed.

Subsequent randomized controlled trials

We did not identify any new eligible RCT.

Table 7. Randomized controlled trials of pulmonary rehabilitation versus conventional care in patients after acute exacerbations of COPD: Trials included in the Puhan et al. systematic review

Study or summary	N_a	Follow-up	Outcome	Findings	Comment
Dyspnea instruments other than CRDQ					
Behnke, 2003	26	18 months	TDI	Between arm difference: • 6.9 (3.9, 9.9) at end of treatment • 8.6 (6.3, 10.9) at 18 months	Exceeds minimal clinically significant difference
Nava, 1997	70	6 weeks	VAS	Between arm difference: • 17 mm ($p<0.01$) favoring PR	Unclear clinical significance
Quality of life					
2 trials (Behnke and Man)	60	12 weeks; 18 months	[a,b]CRDQ dyspnea	Between arm difference: • 1.67 (0.36, 1.98), pHet=0.01, I^2=85%	Very heterogeneous estimate.
2 trials (Behnke and Man)	60	12 weeks; 18 months	[b]CRDQ-fatigue	WMD between arms: • 1.37 (1.13, 1.61), pHet=0.26, I^2=23%	Exceeds minimal clinically significant difference
2 trials (Behnke and Man)	60	12 weeks; 18 months	[b]CRDQ-mastery	WMD between arms: • 1.88 (1.67, 2.09), pHet=0.41, I^2=0%	Exceeds minimal clinically significant difference
2 trials (Behnke and Man)	60	12 weeks; 18 months	[b]CRDQ-function	WMD between arms: • 1.36 (0.94, 1.77), pHet=0.29, I^2=12%	Exceeds minimal clinically significant difference
2 trials (Murphy and Man)	60	12 weeks; 18 months	[c]SGRQ-total	WMD between arms: • -11.1 (-17.1, -5.2), pHet=0.53, I^2=0%	Exceeds minimal clinically significant difference
Functional exercise capacity					
(Behnke, Nava, Troosters)	139	5 weeks to 6 months	[a]6MWT	Between arm difference: • 114 m (28, 199), pHet<0.01, I^2=90%	Very heterogeneous estimate, unclear clinical significance
2 trials (Murphy and Man)	52	NA	SWT	WMD between arms: • 81m (48, 115), pHet=0.55, I^2=0%	Unclear clinical significance
Mortality					
2 trials (Murphy and Man)	52	NA	Mortality	Relative risk: • 0.40 (0.18, 0.86), pHet=0.80, I^2=0%	Including the trial by Nava: relative risk becomes 0.59 (0.34, 1.05), pHet = 0.54, I^2=0%
Hospital admissions					
3 trials (Behnke, Murphy and Man)	93	NA	Hospital admissions	Relative risk: • 0.26 (0.12, 0.54), pHet=0.71, I^2=0%	

CRDQ: Chronic respiratory disease questionnaire; N_a: number analyzed; NA: not applicable; pHet: p-value for heterogeneity; SGRQ: St George's respiratory questionnaire; SWT: shuttle walk test; TDI: transitional dyspnea index; VAS: Visual analog scale; WMD: weighted mean difference
In parentheses in the findings column 95% confidence intervals are shown.
[a] The results were heterogeneous and the Puhan review did not synthesize them.
[b] The review reports that the results are presented on a seven-point scale for CRDQ scores; In this scale the minimal clinically significance difference is 0.5 units.
[c] Negative SGRQ scores favor the PR arm. Minimal clinically significant difference in the total score is 4 units. The minimal clinically significant difference for TDI is 1 unit.

Patients with non-COPD lung disorders

Stable patients

There was only one eligible RCT on participants with non-COPD respiratory disorders especially applicable to the Medicare population.

Asthma

We identified a systematic review assessing the effectiveness and safety of exercise training in asthma[57] that included trials on participants younger than 40 years old. We excluded it because of the age criterion. We did not identify any new eligible RCT.

Bronchiectasis

We excluded one systematic review on physical training for bronchiectasis[58] that included RCT published only as abstracts. One moderate quality (grade B) RCT by Newall et al.[59] compared physical training versus usual care among patients with bronchiectasis. This was a three arm single-center trial that compared endurance exercise training at 80% of peak HR and IMT training (comprehensive PR, n=11), endurance training and sham IMT training (n=12), and usual care (n=9) as a control. The intervention was administered at home/outpatient based PR administered thrice weekly for 8 weeks. Patients were followed up for 3 months.

The comprehensive PR arm had statistically significantly better improvement in the total SGRQ score compared with the control arm both at the end of the intervention (-7.7 [95% confidence interval: -16.6 to 1.1]) and at the end of the follow-up (-10.0 [95% confidence interval: -21.3 to 1.3]).

For maximal exercise capacity (assessed by the ISWT), the differences in the average change from baseline were 113 meters (95% confidence interval: 46, 181) between the first arm and the control at the end of training, but no data were available at end of follow-up.

Functional exercise capacity assessed with a treadmill endurance test improved more in the comprehensive PR arm (720 meters [95% confidence interval: 280, 1160]; digitized figure data) at the end of training.

Three patients experienced acute disease exacerbations by the end of follow-up in the comprehensive PR. The authors reported no safety data.

Patients weaning from mechanical ventilation

We identified no published systematic review for the effects of PR among patients weaning from mechanical ventilation. A single RCT by Porta et al.[60] examined the effects of early arm exercise training on the maximal and functional exercise capacity of the arms. The trial was conducted in three respiratory intensive care units in Italy, among 66 critically ill patients who had successfully weaned from mechanical ventilation (MV) between 48 and 96 hours before trial enrollment. Forty-six (70%) of the patients had COPD; the remaining had restrictive chest wall disease (n=10), "cardiosurgical" sequelae (n=6), septic sequelae (n=2), thoracic trauma (n=1), and abdominal surgery sequel (n=1). The majority (83%) had tracheostomy and almost half (n=31) were on long term O_2 therapy.

Both groups received the standard general physiotherapy treatment with 15 sessions of 20 minutes of additional upper limb exercise in the intervention arm. The exact duration of the program and the length of the follow-up were unclear. Upper limb training increased the maximal exercise capacity of the patients in the arm ergometer (difference in the mean change from baseline between arms 5 W [95% confidence interval: 2, 8] favoring PR); as well as their

endurance time at the constant work rate test (difference in the mean change from baseline 4.1 minutes [95% confidence interval: 0.7, 7.6], favoring PR). The clinical significance of these findings is unclear. The patients' perception of dyspnea in the 10-point modified Borg scale immediately after the tests did not seem to be affected by the intervention.

Table 8 summarizes all the findings of the previous analyses.

Table 8. Summary of the efficacy and safety of pulmonary rehabilitation versus conventional care in different patient populations.

Studies (Patients); Quality ***Effect size (95% confidence interval)***				
Dyspnea and disease-specific QoL		***Maximal exercise capacity***	***Functional exercise capacity***	***Mortality***
Patients with stable COPD – Cochrane Review				
Cochrane review 9 (514) [a]; A|B|C=1|5|3[a]		Cochrane review 12 (464); A|B|C=0|9|3	Cochrane review 11 (486); A|B|C=0|9|2	Cochrane review 7 (542); A|B|C=0|9|2
CRDQ • dyspnea: • fatigue: • mastery: • emotion:	 1.0 (0.8, 1.2) 0.9 (0.7, 1.1) 0.9 (0.7, 1.2) 0.7 (0.4, 1.0)	ICET: 7.1W (2.5, 11.8)[b]	6MWT: 144m (74, 213)	OR = 1.03 (0.54, 1.89)
Patients with stable COPD – RCT published after the Cochrane Review				
Subsequent RCT 1 (40); A|B|C=|0|0|1		ND	Subsequent RCT 2 (246); A|B|C=|1|0|1	ND
CRDQ • dyspnea: • fatigue: • mastery: • emotion:	 0.9 (0.3, 1.5) 0.8 (-0.0, 1.7) 0.8 (-0.1, 1.8) 0.8 (0.0, 1.5)		• 1 RCT with 55 patients: 6MWT: 51m (20, 81)[c] • 1 RCT with 191 patients: 6MWT: NS differences	
Patients after acute exacerbations of COPD				
2 (60); A|B|C=0|1|1		ND	5 (191); A|B|C=0|1|4	3 (122); A|B|C=0|1|2
CRDQ • dyspnea: • fatigue: • mastery: • emotion: SGRQ • total:	 1.7 (0.4, 2.0) 1.4 (1.1, 1.6) 1.9 (1.7, 2.1) 1.4 (0.9, 1.8) -11 (-17, -5)		• 3 RCT with 139 patients: 6MWT: 114m (28, 199) • 2 RCT with 52 patients: SWT:	• 2 RCT with 52 patients (excluding Nava et al.): OR = 0.40 (0.18, 0.86) • 3 RCT with 122 patients (including Nava et al.): OR = 0.59 (0.34, 1.05)
Patients with non-COPD lung disorders - Bronchiectasis				
1 (20); A|B|C=0|1|0		1 (20); A|B|C=0|1|0	1 (20); A|B|C=0|1|0	ND
SGRQ: • total score: Significantly favoring the PR arm		ISWT: 113m (46, 181)	Treadmill endurance test: 720m (280, 1160)	
Patients with non-COPD lung disorders – Weaning from ventilator dependency				
ND		1 (66); A|B|C=0|1|0 Arm ergometry: 5W (2, 8)	1 (66); A|B|C=0|1|0 Endurance time in constant work rate test: 4.1min (0.7, 7.6)	ND

[a] The fatigue, mastery and emotion outcomes were based on 8 studies with 505 patients (a study of 14 patients and B quality did not provide these outcomes) - see Table 2

[b] Some of the primary studies reported kp*m instead of Watts; they have been translated into Watts for these analyses. Two more trials have been excluded from the summary estimates for reasons reported in the text.

[c] At 12 weeks; similar effects but non-significant at 24 weeks (effect was 53m)

6MWT: 6 minute walk test; COPD: chronic obstructive pulmonary disease; CRDQ: chronic respiratory disease questionnaire; ICET: incremental cycle ergometry; ISWT: incremental shuttle walk test; m: meters; min: minutes; NS: not statistically significant; PR: pulmonary rehabilitation; QoL: quality of life; RCT: randomized controlled trial; SWT: shuttle walk test; W: watt(s)

None of these RCT directly assessed the safety of pulmonary rehabilitation interventions. Effect sizes are differences in the change from baseline for single RCT or weighted mean differences thereof for meta-analyses.

Subquestion 1.1: Long term effects of pulmonary rehabilitation

The long term safety of PR has been addressed in the previous sections. We identified a systematic review that assessed the long term outcomes of PR,[21] which was not preferred over the Cochrane review, as already mentioned in the pertinent section. We identified six trials of poor to good quality that potentially assessed the long term efficacy of PR versus usual community care; they were described in 8 papers.[27;30;34;38;39;41;44;61] One did not present any long term outcomes and was excluded from this analysis.[41] Another[44;61] used follow-up interventions in the experimental arm and is described in a separate section. The remaining four assessed long term efficacy outcomes at 12 to 24 months of follow-up. All trials had high attrition rates, but two of them[27;34] analyzed all patients using each participant's latest non-missing measurements. Overall, the effects of the intervention dissipated over time, compared with assessments immediately after the intervention (Table 9). Strijbos[38;39] found no statistically significant differences in the Borg scale for exertional dyspnea. Three trials assessed the CRDQ domains or total score and found statistically significant differences that persisted in the long term (Table 8). Summary CRDQ estimates among the two trials that reported them per domain were 0.5 (95% confidence interval: -0.2, 1.2), 0.7 (95% confidence interval: -0.1, 1.4), 0.7 (95% confidence interval: 0.2, 1.1), and 0.5 (95% confidence interval: 0.1, 0.9) for the dyspnea, fatigue, emotional function and mastery of breath domains, respectively. Thus, the clinical significance of these findings is unclear. Functional exercise capacity in the 6MWT was 93 meters (95% confidence interval: 63, 124) in two trials,[30;34] implying a clinically significant effect among patients who were successfully assessed in the long term. No improvements were documented for the 4MWT and the shuttle walk test in the other two trials. Finally, maximal exercise capacity was not

significantly different across arms in the long term in the two trials that assessed incremental cycle ergometry (summary estimates not available).

Table 9. Randomized controlled trials of pulmonary rehabilitation versus usual community care: Comparison of short term and long term efficacy outcomes (at 1 year or later)

Study, year	***Short term efficacy***		***Long term efficacy***		***Quality***
	Follow-up, N_a	***Findings***	***Follow-up, N_a***	***Findings***	
Griffiths, 2000	6 wk, 184	• CRDQ dyspnea: 1.2 (0.9, 1.5) fatigue: 1.1 (0.7, 1.5) emotion: 1.2 (0.8, 1.5) mastery: 1.1 (0.7, 1.4) • SGRQ, total: -4.7 (-8.5, -0.9) • Functional exercise capacity, SWT: 73m (44, 102)	12 mo, 184	• CRDQ dyspnea: 0.2 (0.1, 0.4) fatigue: 0.3 (-0.1, 0.6) emotion: 0.5 (0.2, 0.8) mastery: 0.4 (-0.0, 0.8) • SGRQ, total: -4 (-8, -0) • Functional exercise capacity, SWT: 16m (-10, 32)	A
Strijbos, 1996	12 wk, 30	• Exertional dyspnea, Borg scale: -0.1 (-3.3, 3.5) • Functional exercise capacity, 4MWT: 31m (-9, 71) • Maximal exercise capacity, ICET: 13W (-1, 26)	18 mo, 27	• Exertional dyspnea, Borg scale: 0.2 (-3.3, 3.7) • Functional exercise capacity, 4MWT: 36m (-9, 81) • Maximal exercise capacity, ICET: 7W (-9, 23)	B
Guell, 1995[a]	18 wk, 56	• CRDQ dyspnea: 1.3 (0.6, 2.0) fatigue: 1.1 (0.5, 1.7) emotion: 1.0 (0.3, 1.7) mastery: 1.2 (0.5, 1.9) • Functional exercise capacity, 6MWT: 83m (48, 118) • Maximal exercise capacity, ICET: 39kp*m (-87, 165)	24 mo, 56	• CRDQ dyspnea: 0.9 (0.4, 1.5) fatigue: 1.1 (0.5, 1.7) emotion: 1.0 (0.3, 1.6) mastery: 0.9 (0.1, 1.8) • Functional exercise capacity, 6MWT: 95m (58, 133) • Maximal exercise capacity, ICET: <20kp*m (p>0.05)	B
Troosters, 2000	24 wk, 62	• CRDQ dyspnea: 0.8 (0.2, 1.5) fatigue: 0.7 (0.1, 1.4) emotion: 0.6 (-0.0, 1.3) mastery: 0.9 (0.2, 1.7) • Functional exercise capacity, 6MWT: 55m (-2, 112) • Maximal exercise capacity, ICET: 11W (-7, 29)	12 mo, 49	• CRDQ, total: 0.9 (0.4, 1.3) (specific domains not available) • Functional exercise capacity, 6MWT: 90m (41, 149) • Maximal exercise capacity, ICET: 13W (4, 22)	B

Trial characteristics have already been presented in Tables 1 to 3. The trial by Booker 1984 also had a long follow-up (12 months) but no results were available. Differences in the change from baseline in the pertinent trials.

[a] Last observation carried forward for 13 dropouts in the long term follow-up

Subquestion 1.2: Relationship between pulmonary rehabilitation-associated harms and comorbid conditions

There was no information on whether comorbid conditions prevalent in the Medicare population affected the risk for PR-associated complications. This is because patients with other comorbid conditions (cardiovascular, orthopedic, metabolic or neurologic diseases) were excluded from PR interventions in the primary trials. In addition reporting of comorbid conditions in the trials was poor.

Subquestion 1.3: Patient level features that modify the effect of pulmonary rehabilitation

Based on the collection of RCT comparing PR versus conventional care, it is unclear what patient level factors affect significantly the effects of PR interventions, either directly, via modifying the participants' gain from the interventions or indirectly, by influencing compliance with the intervention. As mentioned above, the efficacy of PR was not statistically significantly dependent on average disease severity in the corresponding meta-regressions.

Subquestion 1.4: Comparison of pulmonary rehabilitation with general versus individually targeted exercise

We did not identify any published systematic review that compared general exercise protocols with targeted exercise specific to each patient's needs. We identified only one good quality eligible RTC comparing generalized with individualized exercise-based PR interventions.[62] This RCT randomized 180 COPD patients with severe COPD (GOLD IV, mean FEV_1<1.0L) in a highly individualized training scheme versus a generalized exercise training

program. The mean age was 68 years and approximately two thirds of the participants were male. In the individualized training scheme patients were trained only with specific exercises depending on which of their daily activities their disease had the greatest perceived impact. Exercise was performed twice weekly for 7 weeks in an out-patient, supervised setting.

The main outcome was the objective measure of activity using proper devices. There were no statistically significant differences in the four CRDQ domains between the compared arms (the point estimates for differences in the change from baseline for dyspnea, fatigue, emotion and mastery were 0.3 [95% confidence interval: -0.2, 0.8], 0.3 [95% confidence interval: -0.2, 0.8], -0.0 [95% confidence interval:-0.5, 0.4] and 0.1 [95% confidence interval: -0.4, 0.6], respectively). It is evident that the confidence intervals excluded the minimal clinically meaningful difference only for the CRDQ emotional function domain. Patients in both arms were statistically significantly improved compared to baseline ($p<0.0001$).

Both groups improved their maximal exercise capacity in the cycle ergometer significantly compared to baseline ($p<0.0001$), but there was no statistically significant difference in the improvement between the two arms: 3.8m (95% confidence interval: -29.1 to 21.5).

Overall 38 patients experienced acute exacerbations of COPD during the trial (15 versus 23 in the individualized versus generalized training arm, respectively), and 6 patients died (3 versus 3 in the two arms respectively). No data on safety were reported.

Subquestion 1.5: Comparison of pulmonary rehabilitation in different settings and of supervised versus unsupervised pulmonary rehabilitation

We did not find a systematic review that assessed the efficacy of PR in different settings or compared supervised with unsupervised PR. However, we identified three eligible relevant RCT described in five papers.[38;39;63-65]

An RCT of poor methodological quality by Puente-Maestu et al.[64;65] compared physiotherapist-supervised exercise training with unsupervised self-monitored exercise among 49 male patients with severe COPD. Patients in both arms exercised on the treadmill 4 times per week for 8 weeks. Scores in all 4 CRDQ domains improved significantly in both arms after the end of the interventions (unclear whether the improvement was clinically significant). There was no difference between the two arms in the CRDQ scores after the end of the trial. Functional exercise capacity, as measured by endurance time in the cycle ergometer at 70% of baseline peak O_2 uptake improved more in the supervised training arm. The difference in the mean improvement in endurance time was 3.8 minutes (95% confidence interval: 0.7, 7.1). No data on safety were reported.

Another RCT by Elliott et al.[63] compared hospital-based (outpatient, supervised) versus community-based PR (also supervised). Patients were in their sixties, half of them were male and had from moderate to severe COPD. The RCT was of poor methodological quality and had a peculiar design that compared three arms with a complex succession of hospital-based and community-based PR protocols. Because of extensive dropouts (63%) the long term results were neither analyzed nor presented. For the short term results, the two arms that received hospital-based PR were contrasted with the arm that received community-based PR. Both groups showed

significant improvements in the total CRDQ score. However, there was no significant difference in the change from baseline between hospital- and community-based PR for this outcome (6.1 [95% confidence interval: -5 to 17]). The corresponding difference for the 6MWT favored the hospital-based group (70 meters [95% confidence interval: 16 to 123]).

A randomized 3-arm trial of poor methodological quality on 45 COPD patients with severe COPD (GOLD III)[38;39] allowed the comparison of hospital- versus home-based PR. There was no statistically significant difference in exertional dyspnea (modified Borg scale) between the two settings both at the end of the intervention and at 18 months of follow-up. If anything, patients in the hospital-based PR arm tended to claim worse ratings (by 0.1 and 0.4 on average by the end of the intervention and the follow-up, respectively). There were no statistically significant differences between the compared settings at any time point, both for functional exercise capacity (as assessed by the 4MWT) and for maximal exercise capacity (as assessed by the incremental cycle ergometry test). Patients in the hospital-based PR arm tended to have greater improvement in these outcomes compared to patients in the home-based PR arm.

Subquestion 1.6: Efficacy of repeated pulmonary rehabilitation programs

Only one RCT evaluated the repeated programs of PR among subjects with moderate COPD and asthma.[66] The trial enrolled 61 subjects and only 50% were available at the follow-up. The subjects in the intervention arm had PR at baseline, one year, and two years while subjects in the control arm had PR at baseline and second year. Dyspnea was assessed by means of the TDI and Borg scales and HRQOL was measured with the SGRQ instrument. Patients in each PR program had a statistically significant improvement from baseline for the aforementioned outcomes. However, there were no significant differences between the compared

arms. The same pattern was true for functional exercise capacity, which was assessed by the 6MWT: there were improvements beyond chance for each arm, but no statistically significant differences across arms.

Subquestion 1.7: Efficacy and safety of long term maintenance interventions for pulmonary rehabilitation effects

We identified four RCT in five publications that evaluated efforts to maintain the effects of PR after the end of the PR interventions (Table 10).[61;67-70] The trials employed and assessed different strategies to maintain the effects of PR: enhanced follow-up versus conventional follow-up, enhanced follow-up versus no active follow-up, and short term PR versus long term PR. Two trials enrolled subjects with severe COPD;[68;70] one trial enrolled subjects with moderate COPD;[61] and the fourth trial enrolled subjects with mild COPD.[67;69] Home based PR programs were utilized in three trials.[61;68;70] One trial in two publications employed PR on out-patient basis.[67;69] A total of 478 subjects were included in these trials. Follow-up ranged from 12 to 24 months. All trials were graded as B or C for their methodological quality.

Table 10. Description of randomized controlled trials that assessed maintenance interventions after pulmonary rehabilitation.

Author, year	***N_r/N_a***	***Follow-up (months)***	***Interventions***	***Assessed outcomes***	***Overall Quality***
Ries, 2003	161/149	12	After the end of PR • Weekly phone calls and monthly supervised reinforcement sessions • Letter suggesting to continue PR and invitation to monthly alumni group meetings	TDI CRDQ 6MWT Health care use Mortality	B
Berry, 2003 Foy, 2001	140/140	18	• Walking and upper body strength training for 18 months • Walking and upper body strength training for 3 months	CRDQ 6MWT Mortality	C
Wijkstra, 1996	45/36	18	This was a 3 arm RCT: after PR: • Two arms received physical therapy for 30 min 2/day for 3 months and thereafter 1/day • The control arm received no intervention	Borg dyspnea score 6MWT Cycle ergometry	C
Brooks, 2000	85/41	12	• Invited to attend monthly 2 hour group sessions led by a physical therapist, and phone calls between sessions • Conventional follow-up: visited the physical therapist every 3 months	CRDQ SGRQ	C

6MWT: 6 minute walk test; CRDQ: chronic respiratory disease questionnaire; $N_{a/r}$: Number analyzed/randomized; PR: pulmonary rehabilitation; SGRQ: ST George's respiratory questionnaire; TDI: transitional dyspnea index.

Dyspnea score

Two trials assessed dyspnea in severe to moderate COPD using the TDI and Borg scores.[61;70] Both identified that dyspnea worsened over time in the intervention and control arms but there were no differences for the comparisons between arms.

Health related quality of life

Three trials evaluated the changes in the CRDQ.[68-70] One trial compared the changes in the four domains of CRDQ,[69] and the second trial compared changes in the total CRDQ.[70] The third trial by Brooks (reference[68]) described changes in both total and individual domains of CRDQ among severe COPD. For the comparisons of the individual domains of the CRDQ, the decline was significant with time in three of the four domains: dyspnea, fatigue, and mastery. In addition Brooks reported changes in SGRQ that worsened with time but identified no effects across groups. Foy concluded favorable and statistically significant improvement with PR of longer duration among subjects with mild COPD compared to baseline scores in the CRDQ; but there were no significant changes for the comparison between the groups – long term PR versus short term PR.

Two studies that evaluated changes in total CRDQ scores among participants with severe COPD concurred in their results.[68;70] Total quality of life scores were worse at 12 months in the intervention and control arms. Brooks found no significant differences between the groups. Ries reported changes in the total CRDQ among severe COPD that showed significant decline in both groups – maintenance care versus enhanced care (–7 versus –10) as well as statistically and clinically significant differences between the groups.

Functional exercise capacity

All four trials evaluated the 6MWT at long term follow-up. There was a significant decline in the distance over time among subjects with severe COPD in both arms of two trials but no significant differences between the arms.[68;70] Among moderate COPD subjects, the non-active follow-up group showed significant decreases in the walking distance both at 12 and 18 months and no significant changes occurred in those with an active follow-up either with once weekly or once monthly follow-up.[61] However at none of the time points did significant differences occur between the three groups evaluated.

Inspiratory capacity

One trial evaluated the long term effects of PR on the inspiratory capacity (at rest) among subjects with moderate COPD.[61] There was a statistically significant decline in IC with time in subjects who did not have an active follow-up. However these changes were not significant when compared to those with an active follow-up either once weekly or once monthly follow-up.

Key Question 2: Assessment of specific components in exercise-based pulmonary rehabilitation interventions

What is the efficacy and safety of specific PR components in exercise-based PR interventions?

As mentioned in the introduction, we have addressed 4 subquestions in this topic.

Subquestion 2.1: Incremental efficacy and safety of exercise training

Trials pertinent to this subquestion follow the comparison scheme below:

Exercise training +non-exercise PR component(s) versus the same non-exercise PR component(s)

We identified three eligible trials on COPD patients, which were described in four papers (Table 11).[32;71-73] All were small (maximum 66 patients in the compared arms). Applicability for all three RCT was high. The average age of included patients was greater than 65 years. Patients in all trials had from moderate to severe COPD (GOLD III to IV). The RCT by Wedzicha was well designed and well conducted.[71;73] Reporting is suboptimal in the other two trials.[32;72] Exercise training focused mainly on endurance training of ambulatory muscles in all RCT. The trial by Wedzicha[71;73] stratified patients in groups with moderate (Medical Research Council, MRC, 3 to 4) and severe (MRC 5) dyspnea. Attendance was home-based in Larson and the severe dyspnea stratum of Wedzicha and on out-patient basis in the other. The non-exercise PR components were education,[71;73] education and psychological intervention (stress management)[32] and IMT training.[72]

Dyspnea and disease-specific quality of life

There was a statistically significant difference between the compared arms in the total CRDQ score after the completion of the intervention in the moderate dyspnea stratum of the Wedzicha RCT. The authors comment that the 95% CI does not exclude the minimal clinically significant difference in the total score (Table 11). They did not find a statistically significant difference between the compared arms in the severe dyspnea stratum. Moreover, the improvement dissipated after a year of follow-up (p=0.11). Larson et al. examined the CRDQ dyspnea and fatigue scores and found no significant differences between the compared arms (Table 11).

SGRQ was assessed only in the Wedzicha trial.[71;73] Immediately after the intervention, the total SGRQ score favored the arm receiving exercise training (non-significant: -5.5 [95% confidence interval: -10.7, 0.02]). This was not found in the severe dyspnea stratum and disappeared at 1 year of follow-up (p=0.27).

Functional exercise capacity

Larson et al did not find statistically significant differences between exercise training plus IMT and IMT alone in functional exercise capacity (-2 W, [95% confidence interval: –19, 15]). In their moderate dyspnea stratum, Wedzicha[71;73] found an improvement beyond chance in the incremental shuttle walk test favoring the arm that received exercise training (104 meters [95% confidence interval: 60, 148]). During follow-up, performance in both arms deteriorated (more evidently for patients who received exercise training). At the end of the follow-up there was still a statistically significant difference, but of much smaller magnitude (68 meters [95% confidence interval: 11, 125]).

Maximal exercise capacity

Maximal exercise capacity was measured in terms of incremental cycle ergometry testing in two trials.[32;72] Emery found no statistically significant difference in the change from baseline (93J [95% confidence interval: -69, 255]). The same was true for Larson (15W [95% confidence interval: -4, 34]) (Table 11).

All cause mortality

In the Wedzicha trial[71;73] among people in the moderate dyspnea stratum one patient died in the exercise arm versus three patients in the comparator arm by the end of the follow-up (1 year). In the severe dyspnea stratum 0 versus 1 patients died in the corresponding arms by the end of the 8 week intervention.

Safety

The RCT pertinent to the specific question did not evaluate adverse events or complications.

Table 11: Randomized controlled trials assessing the incremental effect of exercise training when added to other non-exercise components.

Study, year	N_r	COPD severity	Male %	Description of compared arms		N_a (%)	Findings	Quality
				Exercise; frequency; setting + Other component	Other component			
Emery, 1998	54	III	54	LLE (End) + ULE; 1/d for 5wk, then 3/wk for 60 to 90 min for 5wk; out-patient + Edu; StMan	Edu; StMan	48 (89)	At end of intervention: • Maximal exercise capacity, [a]ICET: 93 J (95% CI: -69, 255)	B
Larson, 1999	?	III	66	LLE (End); 5d /wk, for 4 months; home-based + IMT	IMT	27 (?)	At end of intervention • CRDQ dyspnea: -0.1 (-1, 0.8) • CRDQ fatigue: 0.5 (-0.3, 1.1) • Functional exercise capacity, CET: -2W (-19, 15) • Maximal exercise capacity, ICET: 15W (-4, 34)	C
Wedzicha, 1998 & Bestall, 2003 (stratum with 3-4 in MRC dyspnea score)	66	IV	44	LLE (End) + ULE; 2/wk for 8 wk; out-patients + Edu After the 8 wks till 12 months, 1/mo and advise to exercise at home between 3 and 5 /wk	Edu	56 (85)	At end of intervention: • [b]ISWT: 104m (60, 148) • [c]CRDQ total 8.9 (2.1, 15.8) • SGRQ total -5.5 (-10.7, 0.02) From baseline to 1 year: • [b]ISWT: 68m (11, 125) • [c]CRDQ total ≅6; p=0.11 • SGRQ total NS, p=0.27 Differences dissipated between end of treatment and end of follow-up in both arms	B
Wedzicha, 1998 (stratum with 5 in MRC dyspnea score)	60	IV	44	LLE (End) + ULE; 2/wk for 8 wk; home-based + Edu	Edu	54 (90)	At end of intervention: • [b]ISWT: -4m (-31, 22) • [c]CRDQ total 0.2 (-4.9, 5.5) • SGRQ total 0.9 (-3.9, 0.8)	B

Trials ordered by COPD severity and then by size.

CET: cycle ergometry test; CRDQ: Chronic respiratory disease questionnaire; Edu: education; End: endurance training; ICET: incremental cycle ergometry test; IMT: inspiratory muscle training; ISWT: incremental shuttle walking test; J: Joule(s); LLE: lower limb exercise; MRC: medical research council; N_a: number analyzed; N_r: number randomized; ULE: upper limb exercise; SGRQ: St George's respiratory questionnaire; StMan: stress management; wk: week(s)

[a] Clinical significance unclear.
[b] this is a maximal exercise capacity outcome
[c] only the total score is given; Note that this is not reported in a 7-point scale where 0.5 units is the minimal clinically significant difference
[e] This is a four-arm parallel trial; male percentage for all arms pooled together

Subquestion 2.2: Efficacy and safety of exercise training compared with other non-exercise PR components

Pertinent trials fit to the comparison scheme described below:

Exercise training versus non-exercise PR component(s)

Five RCT, published between 1985 and 1999 could be used to address this subquestion.[35;72;74-76] All were on COPD patients with moderate to severe disease (GOLD III to IV). Participants were mostly males and aged above 65 years on average. PR was on out-patient basis in all trials. Hence, applicability to the US healthcare system is considered high. Exercise training included endurance training of ambulatory muscles in all but one trial (Baudolff used only upper limb exercise[74]). All RCT were small (maximum was 27 patients in both arms) and overall, reporting of methodological quality items was suboptimal.

Exercise training versus inspiratory muscle training

IMT was the comparator in three RCT.[35;72;76] Only Larson assessed quality of life. They found no differences beyond chance between the compared interventions for the dyspnea and fatigue domains that are reported (point estimates were –0.2 and 0.4 respectively, for the exercise training versus IMT comparison) (Table 12).

Larson and Jones[35;72] assessed maximal exercise capacity by cycle ergometry at the end of the intervention and found no differences beyond chance (point estimates were difference in power of 12 W, and in work of 422 J in the mean changes from baseline, favoring the exercise training arm) (Table 12).

All three assessed functional exercise capacity at end of treatment using either cycle ergometry (Larson[72]) or the 12MWT.[35;76] No significant differences were found between arms (Table 12).

The trials did not assess any survival or safety outcomes.

Exercise training versus education

The RCT by Larson[72] could provide information for this contrast (it is an RCT with four arms). The exercise training arm had a higher improvement in the CRDQ dyspnea score compared with education (difference in the change from baseline of 0.7). However, this was marginally statistically significant and the magnitude of the difference is of unclear clinical significance. No statistically significant differences were found for the CRDQ fatigue domain (the corresponding point estimate was 0.2), the maximal exercise capacity and the functional exercise capacity.

Exercise training versus breathing exercises

The RCT by Berry[75] compared exercise training of ambulatory muscles with breathing exercises. They did not find any statistically significant differences for the mean changes in the modified Borg dyspnea score (mean improvement was 2.5 versus 2.9 in the two arms, respectively). However, they found a significant difference favoring the exercise training arm for the 12MWT functional exercise capacity outcome (p=0.03). No other outcomes of interest were assessed. One patient in each arm had an acute exacerbation of COPD during the intervention.

Exercise training versus phone follow-up

Finally the RCT by Baudolff compared upper limb exercise with phone follow-up.[74] The comparator intervention was perceived as a means to give the same attention to the control group. The only outcome of interest had to do with the functional capacity of the upper arms and was assessed using a ring-moving test. In this test patients are required to move rings passed through a wire for 6 minutes, without touching the wire with the rings. There was a significant improvement in the number of rings moved by the intervention arm compared to baseline ($p=0.03$), but no statistically significant differences were found between treatment arms ($p=0.95$). The clinical significance of this finding is unclear. No safety outcomes were assessed.

Table 12: Randomized controlled trials comparing exercise training with non-exercise components.

Study, year	Size and applicability			Description of compared arms		N_a (%)	Differences between arms	Quality
	N_r	COPD severity	Males (%)	Exercise; frequency; setting	Non-exercise component			
Larson, 1999	?	III	66	LLE (End); 5/wk for 4 months; home-based	Education	26 (?)	At end of intervention • CRDQ dyspnea: 0.7 (+0.0, 1.4) • CRDQ fatigue: 0.2 (-0.7, 1.1) • Functional exercise capacity, CET: -3W (-23, 17) • Maximal exercise capacity, ICET: 15W (-7, 37)	C
Berry, 1996	18	III[a]	64	LLE (End) + ULE (Str); 3/wk for 12 wk; out-patients	Breathing exercises	17 (94)	At end of intervention • Borg dyspnea: mean change from baseline 2.5 vs 2.9, NS • Functional exercise capacity, 12MWT: Significant difference, ANCOVA p=0.03	C
Larson, 1999	?	III	66	LLE (End); 5/wk for 4 months; home-based	IMT	27 (?)	At end of intervention • CRDQ dyspnea: -0.2 (-0.7, 1.0) • CRDQ fatigue: 0.4 (-0.5, 1.2) • Functional exercise capacity, CET: -5W (-21, 12) • Maximal exercise capacity, ICET: 15W (-4, 34)	C
Jones, 1985	22	IV	73	LLE (End) + ULE; for 10 wk; home-based	IMT	15 (68)	At end of intervention • Functional exercise capacity, 12MWT: 36m (-75, 147) • Maximal exercise capacity, ICET: 420 J (-637, 1480)	B
Ries, 1986	18	IV	75	LLE[b] (End); 3/d for 6 wk; home-based	IMT	12 (66)	At end of intervention • Functional exercise capacity, 12MWT: -69m (-326, 188) [b]Endurance time: 4.2min (-3.6, 12.0)	C
Bauldoff, 1996	20	IV	45	ULE (Unsupported, End + Str); 5/wk for 8 wk; home-based	Phone-call follow-up	20 (100)	At end of intervention • Functional exercise capacity, moving of rings Intervention group moved more rings compared to baseline (p=0.03) but no significant differences between arms (p=0.95)	C

Listed by comparator and then by COPD severity and size
12MWT: 12 minute walk test; CET: cycle ergometry test; CRDQ: Chronic respiratory disease questionnaire; End: endurance training; ICET: incremental cycle ergometry test; IMT: inspiratory muscle training; ISWT: incremental shuttle walking test; J: Joule(s) LLE: lower limb exercise; MRC: medical research council; N_a: number analyzed; N_r: number randomized; NS: not significant; ULE: upper limb exercise; SGRQ: St George's respiratory questionnaire; wk: week(s)
[a] baseline values not reported; the estimate is based on post intervention FEV_1 which was 45% and 48% in the 2 arms respectively.
[b] at a rate sustainable for 5 minutes at baseline test

Subquestion 2.3: Incremental efficacy and safety of non-exercise pulmonary rehabilitation components

Pertinent trials follow the comparison scheme described below:

Exercise-based PR +non-exercise PR component(s) versus the same exercise-based PR

We excluded a systematic overview that assessed this question because it did dot provide details on quality of the pertinent RCT.[8]

We identified ten relevant trials on patients with stable COPD,[72;75;77-84] and one trial on patients with bronchiectasis.[59] All were published between 1989 and 2005.

COPD patients

The severity of COPD varied from moderate to severe disease (II to IV in the GOLD classification scheme) (Table 13). In general, participant demographics would be considered analogous to COPD patients in the Medicare patient population. Overall, the trials were very small (minimum 11 and maximum 42 analyzed patients) and overall quality was poor to moderate (usual problems were small sample sizes, suboptimal reporting of methodological quality items and high attrition). Endurance training of ambulatory muscles was included in all exercise-training schemes. The added non-exercise PR component was IMT in nine comparisons, and activity training and lecture series in two comparisons described in a single, three-arm RCT.[81] Administration of PR was inpatient in three RCT.[78;79;82]

Overall, there were no significant differences in the health-related quality of life measures, the functional exercise capacity or the maximal exercise capacity between exercise-based PR without and with the added components. Random effects meta-analyses yielded non-significant differences between the two arms regarding the CRDQ dyspnea, CRDQ fatigue,

maximal and functional exercise capacity and endurance time at 60% to 70% of maximum workload (Table 14). We caution that these results should not be interpreted as proof of equivalence of the compared arms, because of the small sample size of the meta-analyzed trials, their methodological shortcomings, and the fact that all these trials were not designed to assess equivalence. There was a worsening in the CRDQ mastery domain when education was added to exercise training in the Norweg[81] RCT (by 0.8 units on average, adjusted for age, baseline values, and repeated measurements). This finding has limited if any clinical meaning.

The aforementioned RCT did not consistently report data on mortality, acute exacerbations, or safety and harms.

Table 13. Incremental efficacy of non-exercise PR components when added to exercise-based PR among COPD patients.

Study, year	N_r	COPD severity	Male %	Description of components		N_a (%)	Findings (PR + added component vs PR)	Quality
				Intervention common in both arms (Exercise; frequency; setting +non-exercise PR components)	Added component			
Dekhuizen, 1991	40	II/III	75	LLE (End) + torso for 2h; 5/wk for 10 wk; outpatient + Education, breathing exercises, relaxation	IMT	40 (100)	• Functional exercise capacity, 12MWT: 69m (-141, 279) • Maximal exercise capacity, ICET: -7W (-14.4, 0.4)	B
Wanke, 1994	60	II/III	54	LLE (End); 4/wk for 8wk; in-patient	IMT	42 (70)	• Maximal exercise capacity, ICET: 9W (-11, 29)	C
Mador, 2005	38	III	ND	LLE (End); for 8wk; in-patient + Education	IMT	29 (76)	• CRDQ dyspnea: -0.6 (-1.3, +0.0) CRDQ fatigue: -0.7 (-1.4, 0.1) CRDQ emotion: -0.2 (-0.9, 0.6) CRDQ mastery: -0.4 (-1.0, 0.2) • Functional exercise capacity, 6MWT: 4m (-70, 78) Endurance CET at 70% of ICET Wmax: -2.4 min (-12.0 to 7.2) • Maximal exercise capacity, ICET: 0W (-15, 15)	C[a,b]
Larson, 1999	?	III	66	LLE (End); 5d /wk, for 4 months; home-based	IMT	27 (?)	• CRDQ dyspnea: -0.3 (-1, 0.4) CRDQ fatigue: 0.2 (-0.6, 0.9) • Functional exercise capacity, CET: 3W (-16, 22) • Maximal exercise capacity, ICET: 0W (-23, 23)	C
McKeon, 1986	18	III	ND	LLE (End); 7/wk for 6 wk; home-based	IMT	18 (100)	• Functional exercise capacity, 12MWT: NS Endurance stair climbing: NS	C[a]
Berry, 1996	18	III	64	LLE (End) + ULE (Str); 3/wk for 12 wk; out-patient	IMT	17 (94)	• Borg dyspnea score was not significantly different 2.4 (1.2) vs. 2.5 (1.2) • Functional exersise capacity, 12MWT: no significant differences	C
Chen, 1985	13	III	54	LLE (End); 3/wk for 4wk; out-patient	IMT	13 (100)	• Maximal exercise capacity, ICET: 1 W (-23, 25)	C

							• Constant work rate at 66% of Wmax, endurance time: 1 min (-4, 2)	
Weiner, 1992	24	IV	42	LLE + ULE (End + Str) for 30 min; 3/wk for 6mo; outpatient	IMT	24 (100)	• Functional exercise capacity, 12MWT: 435m (173, 697) CET, endurance at 60% of Wmax: 1.9min (-0.8, 4.6)	C
Goldstein, 1989	12	IV	84	LLE + ULE (End); unclear intensity and frequency; in-patient + Education, breathing retraining, relaxation classes	IMT	11 (92)	• Functional exercise capacity, 6MWT: -64m (-198, 70) Submaximal endurance in CET: -1.1 min (-7.1, 4.9)	C[a]
Norweg, 2005[c]	28	II	51	LLE + ULE; 2/wk for 3 wk (supervised) and 2 to 3/wk at home; out-patient + home-based	Activity training	[e]28 (100)	• CRDQ dyspnea 0.7 (p>0.05) CRDQ fatigue 0.6 (p>0.05) CRDQ mastery -0.1 (p>0.05) CRDQ emotion -0.3 (p>0.05) • Functional exercise capacity, 6MWT: no significant difference	B
Norweg, 2005[c]	33	II	51	LLE + ULE; 2/wk for 3 wk (supervised) and 2 to 3/wk at home; out-patient + home-based	Education	[e]33 (100)	• CRDQ dyspnea -0.3 (p>0.05) CRDQ fatigue 0.1 (p>0.05) CRDQ mastery -0.8 (p<0.05) CRDQ emotion -0.7 (p>0.05) • Functional exercise capacity, 6MWT: no significant difference	B

Trials are ordered by added component (IMT, activity training, education), then by COPD severity and decreasing sample size. Main reason for rating the trials as quality C was poor reporting of methodological quality items. Additional reasons are presented when applicable.
A random effects meta-analysis of the available ICET outcomes for the five trials with IMT as the added non-exercise component yields
6/12MWT: 6/12 minute walk test; CET: cycle ergometry test; CRDQ: Chronic respiratory disease questionnaire; End: endurance training; ICET: incremental cycle ergometry test; IMT: inspiratory muscle training; LLE: lower limb exercise; MRC: medical research council; N_a: number analyzed; N_r: number randomized; NS: not significant; ULE: upper limb exercise; wk: week(s)
[a] Suboptimal analysis
[b] Cluster randomized, with unclear description of cluster formation and exact cluster size
[c] This trial has three arms: exercise training alone (n=18), exercise training and activity training (n=10), and exercise training and education (lecture series) (n=15). Optimally analyzed with mixed models accounting for missing information across repeated measurements but with few patients per arm.

Table 14. Random effects meta-analyses of the incremental efficacy of IMT when added to exercise-based PR among COPD patients.

Outcome	**Meta-analysed RCT**	**Number of participants**	**Weighted mean difference (95% confidence interval)**	**Heterogeneity p (I^2 [%])**
CRDQ, dyspnea	Mador, 2005 Larson, 1999	56	-0.5 (-0.9, +0.0)	0.54 (0)
CRDQ, fatigue	Mador, 2005 Larson, 1999	56	-0.3 (-1.1, 0.6)	0.10 (64)
Functional exercise capacity, 12MWT	Dekhuizen, 1991 Weiner, 1992	64	243 m (-115, 601)	0.03 (78)
Maximal exercise capacity, ICET	Dekhuizen, 1991 Wanke, 1994 Mador, 2005 Larson, 1999 Chen, 1985	151	-3.6 W (-9.5, 2.3)	0.59 (0)
Endurance time, constant work rate between 60 and 70% of maximum workload	Mador, 2005 Weiner, 1992 Goldstein, 1989	64	1.3 min (-0.6, 3.3)	0.67 (0)

Trials are ordered as in table 13.

Patients with bronchiectasis

We identified a single RCT by Newall[59] that assessed the incremental impact of IMT in patients with bronchiectasis. Participants attended out-patient, lower limb, exercise training. This was a three-arm trial; here we utilize information from two arms only (exercise training with IMT versus exercise training alone). Generalizability to the Medicare population was good. The RCT was of moderate methodological quality (B in an A to C scale).

There was a trend for better scores in the total SGRQ scale favoring the arm with combined exercise and IMT training at 3 months of follow-up. However, this was statistically non-significant (mean difference in the changes from baseline was –12.3 [95% confidence interval: -24.7, 0.1]). There was no statistically significant difference in the maximal exercise capacity, as measured by the incremental shuttle walk test (difference in the mean changes from baseline was approximately 76 m at 3 months of follow-up, but statistically non-significant). Patients in the combined IMT and exercise arm had greater endurance in a constant work rate treadmill test at 85% of peak O_2 uptake (the corresponding difference was approximately 500 meters).

Subquestion 2.4: Efficacy and safety of different modes of exercise training

Here we assess different exercise training protocols:

a. Higher versus lower intensity training
b. Endurance versus strength exercise training
c. Interval versus continuous training

We identified a published systematic review by Puhan[9] that addressed these three specific questions. Another systematic review by O'Shea et al was identified,[85] but was not selected as

the basis for our assessments. This is because the O'Shea systematic review did not organize the available RCT in meaningful contrasts allowing for a quantitative assessment of the key questions. Thus, it did not provide quantitative syntheses for the outcomes of interest. Moreover, the Puhan[9] systematic review has identified more RCT.

Higher versus lower intensity training

The systematic review identified two trials reported in three papers[86-88] that directly compared higher versus lower intensity training and reported eligible outcomes. Both trials were not eligible according to our inclusion criteria; one did not report any outcome of interest,[88] and the other was on patients younger than 59 years on average.[86;87] We did not identify any other eligible RCT addressing higher versus lower intensity training.

Endurance versus strength training

There were eight relevant RCT reports identified by the systematic review.[79;89-95] We did not identify any other eligible RCT. We present in different sections the comparison between endurance versus strength training (Tables 15 and 16) and the comparison between endurance and combined endurance and strength training (Tables 17 to 19).

There were four RCT that compared endurance versus strength training in COPD patients (disease severity GOLD II to III),[90;91;94;95] all published in 2001 and 2002. Patient characteristics are generalizable to the Medicare population of interest. Reporting of methodological quality items these trials was suboptimal, and two[90;94] had almost 30% attrition rate. One trial was published in German.[95]

Dyspnea outcomes were assessed in three[90;91;94] (Tables 15 and 16) of the four trials. Normandin (n=40) and Ortega (n=33) did not find a difference between the compared exercise modalities in the TDI/BDI instrument (Table 15). All three RCT assessed the four CRDQ

domains (n=103 randomized patients). The changes from baseline in the dyspnea, fatigue and mastery CRDQ domains did not differ beyond chance between the two exercise modalities (Table 15). There was a formally significant difference in favor of strength training for the emotional function domain, but this was only marginally statistically significant and of unclear clinical significance (point estimate -0.4 [95% confidence interval: -0.7, -0.0], Table 15).

Two papers reported data on the 6MWT.[94;95] One of them[95] was published in German and reported outcomes in separate strata of patients (patients with or without O_2 desaturation during training). The overall synthesis showed clinically negligible and non-significant differences between the different exercise modalities (15 meters [95% confidence interval: -14, 44]). The summary synthesis excluded the minimal clinically significant difference of 54 meters in the 6MWT. Results were very consistent across these two trials (Table 16). Another two trials[90;91] showed significantly greater improvements in favor of strength training for constant work rate test endurance (by approximately by 6 minutes in Normandin and 26 minutes in Ortega). However, individual estimates were very different and their confidence intervals did not overlap. Two trials assessed maximal exercise capacity with the incremental cycle ergometry test.[91;94] Their synthesis did not show any formally significant differences between endurance and strength training (4 W [95% confidence interval: –3, 10] favoring endurance exercise training).

There was no difference in the rate of acute COPD exacerbations between the compared arms in the two trials that reported them[90;91] (Table 14, OR 0.67 [95% confidence interval: 0.24, 1.86] favoring endurance training). None of these trials assessed any mortality or safety outcomes.

Table 15. Randomized controlled trials comparing endurance versus strength exercise training, included in the Puhan et al. meta-analysis: health related quality of life (CRDQ) outcomes.

Study, year	N_r	COPD severity	Male %	Exercise; frequency + other components		CRDQ (95% confidence interval), on a seven point scale					Quality
				Endurance	Strength	N_a (%)	Dyspnea	Fatigue	Mastery	*Emotional function*	
Normandin, 2002	54	II/III	53	LLE at 80% of W_{max} for 10-30 min; 3/wk for 10wk + Edu	LLE + ULE, 8-10 repetitions, low intensity; 3/wk for 10 wk + Edu	40 (74)	-0.3 (-0.9, 0.3)	-0.3 (-0.9, 0.3)	-0.2 (-0.6, 0.2)	-0.4 (-1.0, 0.2)	C
Spruit, 2002	48	III	80	LLE 30-75% of W_{max} + ULE for 25 to 60 min; 3/wk for 12 wk	LLE + ULE, 3*8 repetitions at ≥70% of 1 RepMax; 3/wk for 12 wk	30 (63)	0.3 (-0.8, 1.3)	0.0 (-1.0, 1.0)	-0.3, (-1.1, 0.6)	0.0 (-0.8, 0.8)	C
Ortega, 2002	36	III	87	LLE at 60% of W_{max} for 40 min; 3/wk for 12wk + Edu	LLE + ULE, 6-8 repetitions at 70-85% of 1 RepMax; 3/wk for 12 wk + Edu	33 (92)	0.0 (-0.6, 0.6)	-0.4 (-1.0, 0.2)	-0.1 (-0.6, 0.4)	-0.7 (-1.2, -0.2)	B
Overall, random effects meta-analysis						103	-0.1 (-0.5, 0.3)	-0.3 (-0.7, 0.2)	-0.2 (-0.5, 0.2)	-0.4 (-0.7, -0.0)	

Trials ordered by COPD severity and then by decreasing sample size.
Heterogeneity was statistically non-significant for all four domains. Heterogeneity p-values were non-significant and I^2 values were 0% for all 4 outcomes.
CRDQ: Chronic respiratory disease questionnaire; CI: confidence interval; Edu: education; LLE: lower limb exercise; N_a: number analyzed; ND: No data available; N_r: number randomized; NS: not significant; PR: pulmonary rehabilitation; Rel: relaxation techniques; RepMax: (1) repetition maximum; ULE: upper limb exercise; wk: week(s); W_{max}: maximum workload.

Table 16. Randomized controlled trials comparing endurance versus strength exercise training, included in the Puhan et al. meta-analysis: outcomes other than those presented in Table 15.

Study, year	N_r	COPD severity	Male %	Exercise training; frequency + other PR		N_a (%)	Dyspnea and exercise capacity	Other	Quality
				Endurance	Strength				
Wurttemberger, 2001[a]	?	II/III	64	LLE at 70% of W_{max} for 20 min; 3/wk for 3wk + Psy; Rel	LLE + ULE, 2-4*20-25 repetitions at 40% of 1 RepMax; 3/wk for 3 wk + Psy; Rel	46 (?)	• Functional exercise capacity, 6MWT: stratum 1: 17m (-46, 80) stratum 2: 14m (-31, 59)	ND	B
Normandin, 2002	54	II/III	53	LLE at 80% of W_{max} for 10-30 min; 3/wk for 10wk + Edu	LLE + ULE, 8-10 repetitions, low intensity; 3/wk for 10 wk + Edu	40 (74)	• Dyspnea, TDI: -0.3 (-1.7, 1.1) • Functional exercise capacity, CWR: -5.7min (-8.4, 3.0)	• Acute exacerbations 3 vs 4	C
Ortega, 2002	36	III	87	LLE at 60% of W_{max} for 40 min; 3/wk for 12wk + Edu	LLE + ULE, 3*8 repetitions at ≥70% of 1 RepMax; 3/wk for 12 wk + Edu	33 (92)	• Dyspnea, BDI: no SS differences • Functional exercise capacity SWT: -70m (-159, 19) CWR: -25.3min (-38, -13) • Maximal exercise capacity, ICET: 6W (-2, 14)	• Acute exacerbations unclear, but <4 in total	B
Spruit, 2002	48	III	80	LLE 30-75% of W_{max} + ULE for 25 to 60 min; 3/wk for 12 wk	LLE + ULE, 3*8 repetitions at ≥70% of 1 RepMax; 3/wk for 12 wk	30 (63)	• Functional exercise capacity, 6MWT: 16m (-32, 64) • Maximal exercise capacity, ICET: -1W (-12, 10)	• Acute exacerbations 5 vs 7	C

Trials ordered by quality and decreasing sample size.
6MWT: 6 minute walk test; BDI: baseline dyspnea index; BE: breathing exercises; CRDQ: Chronic respiratory disease questionnaire; CWR: constant work rate; Edu: education; LLE: lower limb exercise; m: meter(s); N_a: number analyzed; ND: No data available; N_r: number randomized; NS: not significant; PR: pulmonary rehabilitation; Rel: relaxation techniques; RepMax: (1) repetition maximum; SWT: shuttle walk test; ULE: upper limb exercise; W: watt(s); wk: week(s); W_{max}: maximum workload.
[a] Trial published in German; results presented separately for two different patient strata (with and without O_2 desaturation during exercise). Very consistent estimates with the Spruit et al. trial in the 6MWT.

Finally, the systematic review identified six RCT that compared endurance training versus combined endurance and exercise training.[79;89;91-93;95] These RCT were published between 1988 and 2004, and two (Ortega,[91] Wurtemberger[95]) were mentioned in the previous paragraphs. Two reports, one by Wurttemberger[95] and one from Sivori[93] were published in German and Spanish respectively. Participants were predominantly male and mean ages were 63 years old or older. All were on COPD patients with severity II or III in the GOLD classification scheme. All trials included exercise training of ambulatory muscles (Tables 17 to 19), except for Ries,[92] where upper limb exercise training was employed. Overall, the methodological quality of these RCT was poor to moderate (grades C to B, respectively).

There were no differences between the compared arms for the four CRDQ domains in three trials (n=93 patients in total).[79;89;91] In fact the 95% confidence intervals of the syntheses practically excluded the clinically significant difference of 0.5 units in all four domains (Table 17).

Functional exercise capacity, as conveyed by the 6MWT was assessed in three trials with 106 analyzed participants.[79;89;95] The summary estimates' confidence intervals excluded the minimal clinically significant difference in the 6MWT (54 meters, Table 18). Inferences were similar after excluding data from the German report.[95]

Similarly, there were no statistically significant differences between the compared arms with respect to the maximal exercise capacity in the incremental cycle ergometry test among 165 analyzed COPD patients (1 W [95% confidence interval: -4, 5], Table 17) (references[79;89;91;93;95]). Inferences were very consistent when the non-English language reports[93;95] were excluded. The trial by Ries[92] assessed maximal exercise capacity of the upper limbs, and found no differences between the compared exercise modalities (Table 19).

Finally, no statistically significant differences were found in the constant work exercise tests in two trials by Mador[79] and Ortega[91] (Table 18). The number of patients with acute exacerbations in each arm was not reported in any of the trials (Table 19). Only Ries[92] reported that a patient dropped out of the trial because of low back pain that was attributed to exercise training.

Table 17. Randomized controlled trials comparing endurance versus combined endurance and strength exercise training, included in the Puhan et al. meta-analysis: health related quality of life (CRDQ) outcomes

Study, year	N_r	COPD severity	Male %	Exercise; frequency + other component		CRDQ (95% confidence interval), on a seven point scale					Quality
				Endurance	Endurance/ Strength	N_a (%)	Dyspnea	Fatigue	Mastery	Emotional function	
Bernard, 1999	45	III	78	LLE at 80% of W_{max} for 30 min; 3/wk for 12wk + BE; Rel	LLE endurance/LLE and ULE strength; 3/wk for 12wk + BE; Rel	36 (80)	0.1 (-0.4, 0.6)	0.2 (-0.3, 0.7)	0.6 (-0.1, 1.0)	0.0 (-0.5, 0.5)	B
Ortega, 2002	36	III	87	LLE at 60% of W_{max} for 40 min; 3/wk for 12wk + Edu	LLE endurance/LLE and ULE strength; 3/wk for 12 wk + Edu	31 (86)	0.1 (-0.3, 0.5)	0.1 (-0.4, 0.6)	-0.5 (-0.9, 0.1)	-0.5 (-0.9, -0.1)	B
Mador, 2004	32	III	ND	LLE at ≥50% of W_{max} for 35 min; 3/wk for 8 wk + Edu	LLE endurance/LLE and ULE strength; 3/wk for 8 wk + Edu	24 (75)	0.2 (-0.3, 0.7)	0.2 (-0.5, 0.8)	-0.2 (-0.8, 0.5)	0.0 (-0.3, 0.4)	C
Overall, random effects meta-analysis						93	0.3 (-0.0, 0.5)	0.2 (-0.1, 0.5)	0.1 (-0.3, 0.5)	-0.2 (-0.5, 0.2)	

Trials ordered by quality and decreasing sample size.

Heterogeneity was statistically non-significant for the dyspnea and fatigue domains, and significant ($p<0.1$) for the mastery and emotional function domains.

BE: breathing exercises; CRDQ: Chronic respiratory disease questionnaire; CI: confidence interval; Edu: education; LLE: lower limb exercise; N_a: number analyzed; N_r: number randomized; NS: not significant; PR: pulmonary rehabilitation; Rel: relaxation techniques; ULE: upper limb exercise; wk: week(s); W_{max}: maximum workload.

Table 18. Randomized controlled trials comparing endurance versus combined endurance and strength exercise training, included in the Puhan et al. meta-analysis: functional and maximal exercise capacity outcomes

Study, year	N_r	COPD severity	Male %	Exercise; frequency + other component		Functional and maximal exercise capacity outcomes (95% confidence interval)			Quality
				Endurance	Endurance/ Strength	Na (%)	*6 minute walk test (m)*	*Incremental cycle ergometry (W)*	
Wurttemberger, 2001[a]	?	II/III	64	LLE at 70% of W_{max} for 20 min; 3/wk for 3wk + Psy; Rel	LLE endurance/LLE and ULE strength; 3/wk for 3 wk + Psy; Rel	24 (?)	-5 (-34, 45)	4 (-7, 16)	B
						22 (?)	1 (-48, 50)	3 (-8, 14)	
Bernard, 1999	45	III	78	LLE at 80% of W_{max} for 30 min; 3/wk for 12wk + BE; Rel	LLE endurance/LLE and ULE strength; 3/wk for 12wk + BE; Rel	36 (80)	-22 (-60, 16)	1 (-7, 9)	B
Ortega, 2002	36	III	87	LLE at 60% of W_{max} for 40 min; 3/wk for 12wk + Edu	LLE endurance/LLE and ULE strength; 3/wk for 12 wk + Edu	31 (86)	NA	6 (-5, 17)	B
Sivori, 1998[b]	?	III	89	LLE at 70% of W_{max}; 3/wk for 8wk	LLE endurance/ULE strength; 3/wk for 8wk	28 (?)	NA	-10 (-22, 3)	C
Mador, 2004	32	III	ND	LLE at ≥50% of W_{max} for 35 min; 3/wk for 8 wk + Edu	LLE endurance/LLE and ULE strength; 3/wk for 8 wk + Edu	24 (75)	-8 (-31, 15)	-3 (-16, 9)	C
Overall, random effects meta-analysis							-7 (-24, 9)	1 (-4, 5)	

Trials ordered by severity, and then by quality and decreasing sample size. Heterogeneity was statistically non-significant for both outcomes.
BE: breathing exercises; CWR: constant work rate; Edu: education; LLE: lower limb exercise; m: meter(s); N_a: number analyzed; N_r: number randomized; NS: not significant; Psy: Psychosocial intervention; Rel: relaxation techniques; ULE: upper limb exercise; W: Watt(s); wk: week(s); W_{max}: maximum workload.
The confidence intervals pertain to between-arm differences in the change from baseline.
[a] Trial published in German; results presented separately for two different patient strata (with and without O_2 desaturation during exercise). Results are essentially the same excluding this trial.
[b] Trial published in Spanish; results are essentially the same excluding this trial.

Table 19. Randomized controlled trials comparing endurance versus combined endurance and strength exercise training, included in the Puhan et al. meta-analysis: outcomes other than those reported in Tables 17 and 18

Study, year	Size, applicability			Exercise training; frequency + other PR		N_a (%)	Findings	Quality
	N_r	COPD severity	M (%)	Endurance	Endurance/ Strength			
Bernard, 1999	45	III	78	LLE at 80% of W_{max} for 30 min; 3/wk for 12wk + BE; Rel	LLE endurance/LLE and ULE strength; 3/wk for 12wk + BE; Rel	36 (80)	• Acute exacerbations 3 patients, unclear in which arm	B
Ortega, 2002	36	III	87	LLE; 3/wk for 8wk	LLE endurance/ULE strength; 3/wk for 12 wk + Edu	31 (86)	• Constant work rate test (min): 10 (-4, 23)	C
Mador, 2004	32	III	ND	LLE at ≥50% of W_{max} for 35 min; 3/wk for 8 wk + Edu	LLE endurance/LLE and ULE strength; 3/wk for 8 wk + Edu	24 (75)	• Constant work rate test (min): 0.3 (-7.5, 8.1)	C
Ries, 1988[a]	30	III	ND	LLE at 60% of W_{max} for 40 min; 3/wk for 12wk + Edu	LLE endurance/ULE strength; 3/wk for 12 wk + Edu	18 (60)	• Functional exercise capacity Arm ergometry endurance time: NS differences • Safety; 1 patient dropped out due to low back pain, attributed to the intervention	C

Trials ordered by quality and then by decreasing sample size.
BE: breathing exercises; Edu: education; LLE: lower limb exercise; M: males; N_a: number analyzed; N_r: number randomized; NS: not significant; Rel: relaxation techniques; ULE: upper limb exercise; wk: week(s); W_{max}: maximum workload.
[a] three arm trial; here we contrast the LLE arm with the proprioceptive neuromuscular facilitation arm which received strength training exercises. The remaining arm received endurance exercises of the upper limb.

Continuous versus interval training

The systematic review assessed three trials published between 1999 and 2002 (references[96-98]) that compared continuous training with interval training in people with COPD. One of them (Kaelin[97]) has been reported as a meeting abstract. We identified an additional eligible RCT[99] published after the systematic review on the same topic. We comment on all four trials in the following paragraphs.

Participants were predominantly males (62% to 100%), had COPD severity III to IV in the GOLD classification scheme and their mean age was in the age range of interest. The trials' methodologic quality was poor to moderate. All were small in terms of sample size (13 to 36 analyzed patients). Exercise training of ambulatory muscles was the main exercise in all trials.

Only one trial[98] assessed the four CRDQ domains, and found no differences between the compared arms. Kaelin[97] was the only trial that assessed functional exercise capacity, and found a statistically non-significant trend favoring the continuous exercise arm (41 meters, [95% confidence interval: -17, 99], based on only 13 patients). Maximal exercise capacity with incremental cycle ergometry testing was assessed in three trials.[96;98;99] None found statistically significant differences (Table 20). A meta-analysis of maximal exercise capacity outcomes was not feasible because of missing data on the uncertainty of the estimates in two of the three trials[96;98] (Table 20).

Table 20. Randomized controlled trials of pulmonary rehabilitation interventions comparing continuous versus interval exercise training.

Study, year	N_r	COPD severity	Male %	Exercise training; frequency + other component		N_a (%)	Differences between exercise training modalities	Quality
				Continuous	Interval			
Included in the Puhan et al systematic review								
Vogiatzis, 2002	45	III	62	LLE up to 70% of W_{max}; 2/wk for 12 wk + Edu; BE; Psy; Rel	LLE up to 140% of W_{max} (30 s) and 45% of W_{max} (30 s) + Edu; BE; Psy; Rel	36 (80)	• CRDQ dyspnea: 0.5 (NS) fatigue: 0.0 (NS) mastery: 0.0 (NS) emotion: 0.2 (NS) • Maximal exercise capacity, ICET: -1W (NS)	B
Coppoolse, 1999	21	III	100	LLE at 60% of W_{max}; 5/wk for 8 wk + Edu	LLE at 90% of W_{max} (1 min) and 45% of W_{max} (2min); 3/wk for 8 wk, (also received continuous at 60% of W_{max} 2/wk for 8wk) + Edu	19 (90)	• Maximal exercise capacity ICET: -5 W (NS)	B
Kaelin, 1999 (Abstract)	19	IV	89	LLE; 3/wk for 6 wk + Edu; BE; Psy; Rel	LLE (active to rest ratio 2:1); + Edu; BE; Psy; Rel	13 (68)	• Functional exercise capacity, 6MWT: 41m (-17, 99)	C
Additional trial, published after the Puhan et al. systematic review								
Vogiatzis, 2005	19	III	62	LLE up to 70% of W_{max}; 2/wk for 12 wk + Edu; BE; Psy; Rel	LLE up to 140% of W_{max} (30 s) and 45% of W_{max} (30 s) + Edu; BE; Psy; Rel	19 (100)	• Maximal exercise capacity, ICET: 1W (-25, 27)	B

Trials ordered by worsening COPD severity and then sample size.
BE: breathing exercises; CRDQ: chronic respiratory disease questionnaire; Edu: education; ICET: incremental cycle ergometry test; LLE: lower limb exercise; M: males; N_a: number analyzed; N_r: number randomized; NS: not significant; Psy: psychosocial intervention; Rel: relaxation techniques; ULE: upper limb exercise; W: Watt(s); wk: week(s); W_{max}: maximum workload.

Overview and conclusions on the efficacy and safety of pulmonary rehabilitation

There is little evidence available on the efficacy and safety of PR on diseases other than COPD. In fact, only two trials on other diseases (a trial on idiopathic bronchiectasis and a trial on patients with a variety of diagnoses who were weaning from mechanical ventilation) were eligible according to the inclusion criteria we employed. Their results were similar to the findings of trials on COPD patients. There is also very limited evidence on safety outcomes. There is insufficient information from randomized evidence to compare different PR components. With notable exceptions[27;45;46] almost all included trials were small and potentially underpowered to detect small changes. The majority of the analyzed trials did not report quality items like the randomization method, efforts to conceal patient allocation, efforts to blind the test assessors to intervention, or power analyses. Almost universally, analyses were not by intention-to-treat. Patient follow-up was very short in most trials, and high attrition rates were often observed. Moreover, several trials reported results in figures only, necessitating electronic digitizing from the printed graph, which unavoidably introduced inaccuracies in the quantitative estimates. Finally, the primary trials routinely did not report correlations between the assessed outcomes. This information might be useful, given that the majority of these outcomes are not independent.

Basing a systematic review on existing published systematic reviews poses several challenges. Individual trials may have to be excluded because of differences in eligibility criteria. Updating may also be complicated because it is often difficult to capture the exact inclusion and exclusion criteria of the primary reviews. Individual systematic reviews use

different RCT quality scoring systems. Despite efforts to standardize, inconsistencies may persist. The analysis approach may also be variable: In the two systematic reviews by Puhan et al. the primary reviewers decided not to combine heterogeneous trials, whereas we performed meta-analyses using random effects models. We added or subtracted trials from the Cochrane review meta-analyses according to our criteria, and we performed different subgroup analyses than those of the Cochrane review. Allowing for these caveats, it is unlikely that the aforementioned challenges would invalidate our results.

Efficacy of exercise-based pulmonary rehabilitation versus conventional care

Existing evidence indicates that exercise-based PR interventions are efficacious in the short term. Fewer trials have assessed long term efficacy outcomes, and their results are in agreement with the short term findings. As noted above, almost all eligible trials pertained to patients with stable COPD or patients after acute exacerbations of COPD.

More specifically, exercise-based PR improves patients' quality of life, maximal and functional exercise capacity beyond what would be expected by chance. Especially in the short term, the improvements in three domains of the CRDQ instrument (namely dyspnea, fatigue and mastery) and in the 6MWT were significantly greater than the minimal clinically significant differences in these outcomes. There is no evidence that the benefits of PR are translated into survival differences, at least among people with stable COPD. This is not surprising, given that few RCT extended follow-up beyond 12 months, and deaths are just too sparse in the short term to detect a statistically significant difference. However, exercise-based PR interventions may reduce hospitalizations and primary care consultations. We believe that the existing RCT suffice to appreciate the the short-term efficacy of exercise-based PR, at least in COPD.

There is also evidence favoring the efficacy of exercise-based PR among patients who are recovering from, or recently recovered from acute exacerbations of COPD. In fact the claimed effects in the health-related quality of life and exercise capacity outcomes are even larger in this patient subgroup. This may be ascribed to the fact that PR accelerates the participants' recovery. An alternative explanation would be that a "hedonistic treadmill" phenomenon has been observed, especially for quality of life outcomes: Health-related quality of life may be very dismally perceived during the acute illness, and an "overcorrection" in this perception might be observed as soon as the overall condition improves. Mortality and hospital re-admissions appear to decrease with exercise-PR based interventions after acute COPD exacerbations. We caution that all these results are from a few small trials with methodological shortcomings, and thus they might be overestimations of the true effects. Larger RCT are needed to confirm these findings.

Safety of exercise-based pulmonary rehabilitation versus conventional care

Data on safety were very sparsely reported. This paucity of data should not be viewed as evidence of absence of adverse events.

Overall, little is known about the harms associated with PR interventions and which comorbid conditions predispose patients to or protect them from these adverse events. It is anticipated that many comorbid conditions are present in older COPD patients who undergo pulmonary rehabilitation. However, the eligible studies provided little information on the presence of such comorbidities in the studied populations. As expected, patients with serious comorbidities that might have affected the ability of the patients to exercise (unstable cardiac disease, orthopedic and musculoskeletal disease, malignancies etc.) were routinely excluded from PR trials.

Relative value of different exercise training protocols and of different pulmonary rehabilitation components

Our analysis focused on comparisons with exercise training, and we did not assess the effects on non-exercise components versus no intervention. Overall, information is limited, and based on trials of very small sample sizes. We should caution that the absence of statistically significant differences does not imply that the compared protocols are equivalent. Sample sizes are just too small and the enrolled RCT were not designed to assess equivalence or non-inferiority.

More specifically, there seems to be no formally significant difference between exercise protocols that are tailored to address each patient's specific weaknesses and exercise protocols that are common for all patients. Similarly, there was no evidence in favor of repeated PR interventions, or additional interventions employed to maintain the effects of PR. Poor quality trials reported that supervised and hospital based PR may be advantageous over unsupervised and community or home-based PR, respectively.

Strength training was not consistently associated beyond chance with more favorable outcomes compared to endurance training in the trials that directly compared the two training modalities. This was true when strength training was compared with combined strength and endurance training in a different set of RCT. Interval training protocols may be another option to the continuous training protocols that are usually employed. Finally, there were no RCT that were applicable to the Medicare population of interest that compared high and low intensity training.

Sparse data suggested that exercise training tended to have an additive impact when added to non-exercise PR components like education and/or psychosocial interventions, at least

for health-related quality of life. Compared with education alone, exercise and education confer additional benefits in health-related quality of life (total CRDQ) and functional exercise capacity in subjects with moderate functional limitation resulting from dyspnea. The clinical significance of the observed differences was unclear, mainly because of small sample sizes. The same was true when exercise-only PR is contrasted with non-exercise interventions (i.e., IMT, education, breathing, phone follow-up). Because of small sample sizes, very few significant differences were observed.

Finally, we did not find statistically significant differences when we assessed combined exercise training and non-exercise components (i.e., IMT, activity training and lecture series) versus exercise training alone. However, these results should be regarded as proof of equivalence between the compared interventions. They should be viewed with caution because of the limited sample sizes and the questionable methodologic validity of many of the included trials.

References

(1) Pulmonary rehabilitation-1999. American Thoracic Society. Am J Respir Crit Care Med 1999; 159(5 Pt 1):1666-1682.

(2) Nici L, Donner C, Wouters E, Zuwallack R, Ambrosino N, Bourbeau J et al. American Thoracic Society/European Respiratory Society statement on pulmonary rehabilitation. Am J Respir Crit Care Med 2006; 173(12):1390-1413.

(3) Pulmonary rehabilitation: joint ACCP/AACVPR evidence-based guidelines. ACCP/AACVPR Pulmonary Rehabilitation Guidelines Panel. American College of Chest Physicians. American Association of Cardiovascular and Pulmonary Rehabilitation. Chest 1997; 112(5):1363-1396.

(4) Pulmonary rehabilitation. Thorax 2001; 56(11):827-834.

(5) Hill NS. Pulmonary rehabilitation. Proc Am Thorac Soc 2006; 3(1):66-74.

(6) Lacasse Y, Wong E, Guyatt GH, King D, Cook DJ, Goldstein RS. Meta-analysis of respiratory rehabilitation in chronic obstructive pulmonary disease.[see comment]. Lancet 1996; 348(9035):1115-1119.

(7) Lacasse Y, Brosseau L, Milne S, Martin S, Wong E, Guyatt GH et al. Pulmonary rehabilitation for chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2002;(3):CD003793.

(8) Lacasse Y, Guyatt GH, Goldstein RS. The components of a respiratory rehabilitation program: a systematic overview.[see comment]. Chest 1997; 111(4):1077-1088.

(9) Puhan MA, Schunemann HJ, Frey M, Scharplatz M, Bachmann LM. How should COPD patients exercise during respiratory rehabilitation? Comparison of exercise modalities and intensities to treat skeletal muscle dysfunction. Thorax 2005; 60(5):367-375.

(10) Puhan MA, Scharplatz M, Troosters T, Steurer J. Respiratory rehabilitation after acute exacerbation of COPD may reduce risk for readmission and mortality -- a systematic review. Respir Res 2005; 6(1):54.

(11) Celli BR, Cote CG, Marin JM, Casanova C, Montes de OM, Mendez RA et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease.[see comment]. New England Journal of Medicine 2004; 350(10):1005-1012.

(12) Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. Am J Respir Crit Care Med 2001; 163(5):1256-1276.

(13) Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. Chest 1988; 93(3):580-586.

(14) Guyatt GH, Berman LB, Townsend M, Pugsley SO, Chambers LW. A measure of quality of life for clinical trials in chronic lung disease. Thorax 1987; 42(10):773-778.

(15) Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. Am Rev Respir Dis 1992; 145(6):1321-1327.

(16) The Handbook of Research Synthesis. New York: Russell Sage Foundation; 1994.

(17) DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7(3):177-188.

(18) Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002; 21(11):1539-1558.

(19) Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. Stat Med 2003; 22(17):2693-2710.

(20) Salman GF, Mosier MC, Beasley BW, Calkins DR. Rehabilitation for patients with chronic obstructive pulmonary disease: meta-analysis of randomized controlled trials.[see comment]. Journal of General Internal Medicine 2003; 18(3):213-221.

(21) Cambach W, Wagenaar RC, Koelman TW, van Keimpema AR, Kemper HC. The long-term effects of pulmonary rehabilitation in patients with asthma and chronic obstructive pulmonary disease: a research synthesis. Archives of Physical Medicine & Rehabilitation 1999; 80(1):103-111.

(22) Chavannes N, Vollenberg JJ, van Schayck CP, Wouters EF. Effects of physical activity in mild to moderate COPD: a systematic review. Br J Gen Pract 2002; 52(480):574-578.

(23) Sin DD, McAlister FA, Man SF, Anthonisen NR. Contemporary management of chronic obstructive pulmonary disease: scientific review. JAMA 2003; 290(17):2301-2312.

(24) Busch AJ, McClements JD. Effects of a supervised home exercise program on patients with severe chronic obstructive pulmonary disease. Physical Therapy 1988; 68(4):469-474.

(25) Cambach W, Chadwick-Straver RV, Wagenaar RC, van Keimpema AR, Kemper HC. The effects of a community-based pulmonary rehabilitation programme on exercise tolerance and quality of life: a randomized controlled trial. European Respiratory Journal 1997; 10(1):104-113.

(26) Goldstein RS, Gort EH, Stubbing D, Avendano MA, Guyatt GH. Randomised controlled trial of respiratory rehabilitation.[see comment]. Lancet 1994; 344(8934):1394-1397.

(27) Griffiths TL, Burr ML, Campbell IA, Lewis-Jenkins V, Mullins J, Shiels K et al. Results at 1 year of outpatient multidisciplinary pulmonary rehabilitation: a randomised controlled trial.[see comment][erratum appears in Lancet 2000 Apr 8;355(9211):1280 Note: Lonescu AA [corrected to Ionescu AA]]. Lancet 2000; 355(9201):362-368.

(28) Hernandez MT, Rubio TM, Ruiz FO, Riera HS, Gil RS, Gomez JC. Results of a home-based training program for patients with COPD. Chest 2000; 118(1):106-114.

(29) Simpson K, Killian K, McCartney N, Stubbing DG, Jones NL. Randomised controlled trial of weightlifting exercise in patients with chronic airflow limitation. Thorax 1992; 47(2):70-75.

(30) Troosters T, Gosselink R, Decramer M. Short- and long-term effects of outpatient rehabilitation in patients with chronic obstructive pulmonary disease: a randomized trial. American Journal of Medicine 2000; 109(3):207-212.

(31) Wijkstra PJ, van AR, Kraan J, Otten V, Postma DS, Koeter GH. Quality of life in patients with chronic obstructive pulmonary disease improves after rehabilitation at home. European Respiratory Journal 1994; 7(2):269-273.

(32) Emery CF, Schein RL, Hauck ER, MacIntyre NR. Psychological and cognitive outcomes of a randomized trial of exercise among patients with chronic obstructive pulmonary disease. Health Psychology 1998; 17(3):232-240.

(33) Engstrom CP, Persson LO, Larsson S, Sullivan M. Long-term effects of a pulmonary rehabilitation programme in outpatients with chronic obstructive pulmonary disease: a randomized controlled study. Scandinavian Journal of Rehabilitation Medicine 1999; 31(4):207-213.

(34) Guell R, Casan P, Belda J, Sangenis M, Morante F, Guyatt GH et al. Long-term effects of outpatient rehabilitation of COPD: A randomized trial. Chest 2000; 117(4):976-983.

(35) Jones DT, Thomson RJ, Sears MR. Physical exercise and resistive breathing training in severe chronic airways obstruction--are they effective? Eur J Respir Dis 1985; 67(3):159-166.

(36) Lake FR, Henderson K, Briffa T, Openshaw J, Musk AW. Upper-limb and lower-limb exercise training in patients with chronic airflow obstruction. Chest 1990; 97(5):1077-1082.

(37) McGavin CR, Gupta SP, Lloyd EL, McHardy GJ. Physical rehabilitation for the chronic bronchitic: results of a controlled trial of exercises in the home. Thorax 1977; 32(3):307-311.

(38) Strijbos JH, Postma DS, van AR, Gimeno F, Koeter GH. A comparison between an outpatient hospital-based pulmonary rehabilitation program and a home-care pulmonary rehabilitation program in patients with COPD. A follow-up of 18 months.[see comment]. Chest 1996; 109(2):366-372.

(39) Strijbos JH, Postma DS, van AR, Gimeno F, Koeter GH. Feasibility and effects of a home-care rehabilitation program in patients with chronic obstructive pulmonary disease. Journal of Cardiopulmonary Rehabilitation 1996; 16(6):386-393.

(40) Vallet G, Varray A, Fontaine JL, Prefaut C. [Value of individualized rehabilitation at the ventilatory threshold level in moderately severe chronic obstructive pulmonary disease]. Rev Mal Respir 1994; 11(5):493-501.

(41) Booker HA. Exercise training and breathing control in patients with chronic airflow limitation. Physiotherapy 1984; 70(7):258-260.

(42) Ringbaek TJ, Broendum E, Hemmingsen L, Lybeck K, Nielsen D, Andersen C et al. Rehabilitation of patients with chronic obstructive pulmonary disease. Exercise twice a week is not sufficient! Respiratory Medicine 2000; 94(2):150-154.

(43) Bendstrup KE, Ingemann JJ, Holm S, Bengtsson B. Out-patient rehabilitation improves activities of daily living, quality of life and exercise tolerance in chronic obstructive pulmonary disease. European Respiratory Journal 1997; 10(12):2801-2806.

(44) Wijkstra PJ, ten VERGERT EM, van AR, Otten V, Kraan J, Postma DS et al. Long term benefits of rehabilitation at home on quality of life and exercise tolerance in patients with chronic obstructive pulmonary disease. Thorax 1995; 50(8):824-828.

(45) Bourbeau J, Julien M, Maltais F, Rouleau M, Beaupre A, Begin R et al. Reduction of hospital utilization in patients with chronic obstructive pulmonary disease: a disease-specific self-management intervention.[see comment]. Archives of Internal Medicine 2003; 163(5):585-591.

(46) Gadoury MA, Schwartzman K, Rouleau M, Maltais F, Julien M, Beaupre A et al. Self-management reduces both short- and long-term hospitalisation in COPD. European Respiratory Journal 2005; 26(5):853-857.

(47) Finnerty JP, Keeping I, Bullough I, Jones J. The effectiveness of outpatient pulmonary rehabilitation in chronic lung disease: a randomized controlled trial. Chest 2001; 119(6):1705-1710.

(48) Singh V, Khandelwal DC, Khandelwal R, Abusaria S. Pulmonary rehabilitation in patients with chronic obstructive pulmonary disease. Indian Journal of Chest Diseases & Allied Sciences 2003; 45(1):13-17.

(49) Kirsten DK, Taube C, Lehnigk B, Jorres RA, Magnussen H. Exercise training improves recovery in patients with COPD after an acute exacerbation. Respiratory Medicine 1998; 92(10):1191-1198.

(50) Behnke M, Taube C, Kirsten D, Lehnigk B, Jorres RA, Magnussen H. Home-based exercise is capable of preserving hospital-based improvements in severe chronic obstructive pulmonary disease. Respiratory Medicine 2000; 94(12):1184-1191.

(51) Behnke M, Jorres RA, Kirsten D, Magnussen H. Clinical benefits of a combined hospital and home-based exercise programme over 18 months in patients with severe COPD. Monaldi Archives for Chest Disease 2003; 59(1):44-51.

(52) Man WD, Polkey MI, Donaldson N, Gray BJ, Moxham J. Community pulmonary rehabilitation after hospitalisation for acute exacerbations of chronic obstructive pulmonary disease: randomised controlled study.[see comment]. BMJ 2004; 329(7476):1209.

(53) Murphy N, Bell C, Costello RW. Extending a home from hospital care programme for COPD exacerbations to include pulmonary rehabilitation. Respir Med 2005; 99(10):1297-1302.

(54) Nava S. Rehabilitation of patients admitted to a respiratory intensive care unit. Archives of Physical Medicine & Rehabilitation 1998; 79(7):849-854.

(55) Troosters T, Gosselink R, De Paepe K. Pulmonary rehabilitation improves survival in COPD patients with a recent severe acute exacerbation. Am J Respir Crit Care Med 2002; 165:A16.

(56) Witek TJ, Jr., Mahler DA. Minimal important difference of the transition dyspnoea index in a multinational clinical trial. Eur Respir J 2003; 21(2):267-272.

(57) Ram FS, Robinson SM, Black PN. Effects of physical training in asthma: a systematic review. Br J Sports Med 2000; 34(3):162-167.

(58) Bradley J, Moran F, Greenstone M. Physical training for bronchiectasis. Cochrane Database Syst Rev 2002;(3):CD002166.

(59) Newall C, Stockley RA, Hill SL. Exercise training and inspiratory muscle training in patients with bronchiectasis.[see comment]. Thorax 2005; 60(11):943-948.

(60) Porta R, Vitacca M, Gile LS, Clini E, Bianchi L, Zanotti E et al. Supported arm training in patients recently weaned from mechanical ventilation. Chest 2005; 128(4):2511-2520.

(61) Wijkstra PJ, van der Mark TW, Kraan J, van AR, Koeter GH, Postma DS. Long-term effects of home rehabilitation on physical performance in chronic obstructive pulmonary disease. American Journal of Respiratory & Critical Care Medicine 1996; 153(4:Pt 1):t-41.

(62) Sewell L, Singh SJ, Williams JE, Collier R, Morgan MD. Can individualized rehabilitation improve functional independence in elderly patients with COPD? Chest 2005; 128(3):1194-1200.

(63) Elliott M, Watson C, Wilkinson E, Musk AW, Lake FR. Short- and long-term hospital and community exercise programmes for patients with chronic obstructive pulmonary disease. Respirology 2004; 9(3):345-351.

(64) Puente-Maestu L, Sanz ML, Sanz P, Cubillo JM, Mayol J, Casaburi R. Comparison of effects of supervised versus self-monitored training programmes in patients with chronic obstructive pulmonary disease. European Respiratory Journal 2000; 15(3):517-525.

(65) Puente-Maestu L, Sanz ML, Sanz P, Ruiz de Ona JM, Rodriguez-Hermosa JL, Whipp BJ. Effects of two types of training on pulmonary and cardiac responses to moderate exercise in patients with COPD. European Respiratory Journal 2000; 15(6):1026-1032.

(66) Foglio K, Bianchi L, Ambrosino N. Is it really useful to repeat outpatient pulmonary rehabilitation programs in patients with chronic airway obstruction? A 2-year controlled study. Chest 2001; 119(6):1696-1704.

(67) Berry MJ, Rejeski WJ, Adair NE, Ettinger WH, Jr., Zaccaro DJ, Sevick MA. A randomized, controlled trial comparing long-term and short-term exercise in patients with chronic obstructive pulmonary disease. Journal of Cardiopulmonary Rehabilitation 2003; 23(1):60-68.

(68) Brooks D, Krip B, Mangovski-Alzamora S, Goldstein RS. The effect of postrehabilitation programmes among individuals with chronic obstructive pulmonary disease.[see comment]. European Respiratory Journal 2002; 20(1):20-29.

(69) Foy CG, Rejeski WJ, Berry MJ, Zaccaro D, Woodard CM. Gender moderates the effects of exercise therapy on health-related quality of life among COPD patients. Chest 2001; 119(1):70-76.

(70) Ries AL, Kaplan RM, Myers R, Prewitt LM. Maintenance after pulmonary rehabilitation in chronic lung disease: a randomized trial. American Journal of Respiratory & Critical Care Medicine 2003; 167(6):880-888.

(71) Bestall JC, Paul EA, Garrod R, Garnham R, Jones RW, Wedzicha AJ. Longitudinal trends in exercise capacity and health status after pulmonary rehabilitation in patients with COPD. Respiratory Medicine 2003; 97(2):173-180.

(72) Larson JL, Covey MK, Wirtz SE, Berry JK, Alex CG, Langbein WE et al. Cycle ergometer and inspiratory muscle training in chronic obstructive pulmonary disease. American Journal of Respiratory & Critical Care Medicine 1999; 160(2):500-507.

(73) Wedzicha JA, Bestall JC, Garrod R, Garnham R, Paul EA, Jones PW. Randomized controlled trial of pulmonary rehabilitation in severe chronic obstructive pulmonary disease patients, stratified with the MRC dyspnoea scale. European Respiratory Journal 1998; 12(2):363-369.

(74) Bauldoff GS, Hoffman LA, Sciurba F, Zullo TG. Home-based, upper-arm exercise training for patients with chronic obstructive pulmonary disease. Heart & Lung 1996; 25(4):288-294.

(75) Berry MJ, Adair NE, Sevensky KS, Quinby A, Lever HM. Inspiratory muscle training and whole-body reconditioning in chronic obstructive pulmonary disease. American Journal of Respiratory & Critical Care Medicine 1996; 153(6:Pt 1):t-6.

(76) Ries AL, Moser KM. Comparison of isocapnic hyperventilation and walking exercise training at home in pulmonary rehabilitation. Chest 1986; 90(2):285-289.

(77) Dekhuijzen PN, Folgering HT, van Herwaarden CL. Target-flow inspiratory muscle training during pulmonary rehabilitation in patients with COPD. Chest 1991; 99(1):128-133.

(78) Goldstein R, De RJ, Long S, Dolmage T, Avendano MA. Applicability of a threshold loading device for inspiratory muscle testing and training in patients with COPD. Chest 1989; 96(3):564-571.

(79) Mador MJ, Deniz O, Aggarwal A, Shaffer M, Kufel TJ, Spengler CM. Effect of respiratory muscle endurance training in patients with COPD undergoing pulmonary rehabilitation. Chest 2005; 128(3):1216-1224.

(80) McKeon JL, Turner J, Kelly C, Dent A, Zimmerman PV. The effect of inspiratory resistive training on exercise capacity in optimally treated patients with severe chronic airflow limitation. Australian & New Zealand Journal of Medicine 1986; 16(5):648-652.

(81) Norweg AM, Whiteson J, Malgady R, Mola A, Rey M. The effectiveness of different combinations of pulmonary rehabilitation program components: a randomized controlled trial. Chest 2005; 128(2):663-672.

(82) Wanke T, Formanek D, Lahrmann H, Brath H, Wild M, Wagner C et al. Effects of combined inspiratory muscle and cycle ergometer training on exercise performance in patients with COPD.[see comment]. European Respiratory Journal 1994; 7(12):2205-2211.

(83) Weiner P, Azgad Y, Ganam R. Inspiratory muscle training combined with general exercise reconditioning in patients with COPD. Chest 1992; 102(5):1351-1356.

(84) Chen H, Dukes R, Martin BJ. Inspiratory muscle training in patients with chronic obstructive pulmonary disease. American Review of Respiratory Disease 1985; 131(2):251-255.

(85) O'Shea SD, Taylor NF, Paratz J. Peripheral muscle strength training in COPD: a systematic review. Chest 2004; 126(3):903-914.

(86) Casaburi R, Patessio A, Ioli F, Zanaboni S, Donner CF, Wasserman K. Reductions in exercise lactic acidosis and ventilation as a result of exercise training in patients with obstructive lung disease. Am Rev Respir Dis 1991; 143(1):9-18.

(87) Patessio A, Carone M, Ioli F, Donner CF. Ventilatory and metabolic changes as a result of exercise training in COPD patients. Chest 1992; 101(5:Suppl):Suppl-278S.

(88) Vallet G, Ahmaidi S, Serres I, Fabre C, Bourgouin D, Desplan J et al. Comparison of two training programmes in chronic airway limitation patients: standardized versus individualized protocols. Eur Respir J 1997; 10(1):114-122.

(89) Bernard S, Whittom F, LeBlanc P, Jobin J, Belleau R, Berube C et al. Aerobic and strength training in patients with chronic obstructive pulmonary disease. American Journal of Respiratory & Critical Care Medicine 1999; 159(3):896-901.

(90) Normandin EA, McCusker C, Connors M, Vale F, Gerardi D, ZuWallack RL. An evaluation of two approaches to exercise conditioning in pulmonary rehabilitation. Chest 2002; 121(4):1085-1091.

(91) Ortega F, Toral J, Cejudo P, Villagomez R, Sanchez H, Castillo J et al. Comparison of effects of strength and endurance training in patients with chronic obstructive pulmonary disease. American Journal of Respiratory & Critical Care Medicine 2002; 166(5):669-674.

(92) Ries AL, Ellis B, Hawkins RW. Upper extremity exercise training in chronic obstructive pulmonary disease. Chest 1988; 93(4):688-692.

(93) Sivori M, Rhodius E, Kaplan P, Talarico M, Gorojod G, Carreras B et al. [Exercise training in chronic obstructive pulmonary disease. Comparative study of aerobic training of lower limbs vs. combination with upper limbs]. Medicina (B Aires) 1998; 58(6):717-727.

(94) Spruit MA, Gosselink R, Troosters T, De PK, Decramer M. Resistance versus endurance training in patients with COPD and peripheral muscle weakness. European Respiratory Journal 2002; 19(6):1072-1078.

(95) Wurtemberger G, Bastian K. [Functional effects of different training in patients with COPD]. Pneumologie 2001; 55(12):553-562.

(96) Coppoolse R, Schols AM, Baarends EM, Mostert R, Akkermans MA, Janssen PP et al. Interval versus continuous training in patients with severe COPD: a randomized clinical trial. European Respiratory Journal 1999; 14(2):258-263.

(97) Results of 6 minute ambulation and met tolerance of patients with severe chronic obstructive pulmonary disease using two different aerobic training regimens: interval training vs. continuous training. American Society of Exercise Physiology 2nd Annual Meeting; 1997.

(98) Vogiatzis I, Nanas S, Roussos C. Interval training as an alternative modality to continuous exercise in patients with COPD.[see comment]. European Respiratory Journal 2002; 20(1):12-19.

(99) Vogiatzis I, Terzis G, Nanas S, Stratakos G, Simoes DC, Georgiadou O et al. Skeletal muscle adaptations to interval training in patients with advanced COPD. Chest 2005; 128(6):3838-3845.

Appendices

There are three appendices to the technology assessment.

A. Detailed search strategy

B. Comments on validated outcomes

C. The Jadad quality scale

D. Evidence tables for RCT not included in the published systematic reviews used in this report

Appendix A. Detailed search strategy

The following search strategy was used (formulated for OVID MEDLINE)

1 exp lung diseases/
2 exp asthma/
3 exp pulmonary disease, chronic obstructive/
4 exp bronchiectasis/
5 exp respiration, artificial/
6 exp lung transplantation/
7 obstructive pulmonary disease$.tw.
8 COPD.tw.
9 or/1-8
10 exp rehabilitation/
11 exp exercise therapy/
12 rehabilit$.mp.
13 exp Exercise Movement Techniques/
14 exp exercise tolerance/
15 exp physical therapy modalities/
16 rh.fs.
17 or/10-16
18 9 and 17
19 follow-up studies/
20 (follow-up or followup).tw.
21 exp Case-Control Studies/
22 (case adj20 control).tw.
23 exp Longitudinal Studies/
24 longitudinal.tw.
25 exp Cohort Studies/
26 cohort.tw.
27 (random$ or rct).tw.
28 exp Randomized Controlled Trials/
29 exp random allocation/
30 exp Double-Blind Method/
31 exp Single-Blind Method/
32 randomized controlled trial.pt.
33 clinical trial.pt.
34 controlled clinical trials/
35 (clin$ adj trial$).tw.
36 ((singl$ or doubl$ or trebl$ or tripl$) adj (blind$ or mask$)).tw.
37 exp PLACEBOS/
38 placebo$.tw.
39 exp Research Design/
40 exp Evaluation Studies/
41 exp Prospective Studies/

42 exp Comparative Study/
43 or/19-42
44 18 and 43
45 limit 44 to humans
46 limit 45 to english language
47 limit 46 to (addresses or bibliography or biography or case reports or congresses or consensus development conference or consensus development conference, nih or dictionary or directory or festschrift or government publications or guideline or interview or lectures or legal cases or legislation or news or newspaper article or patient education handout or periodical index or practice guideline or "review")
48 46 not 47
49 limit 46 to (editorial and comment)
50 limit 46 to (letter and comment)
51 limit 46 to (editorial and letter)
52 or/49-51
53 limit 52 to clinical trial
54 52 not 53
55 46 not (47 or 54)
56 limit 55 to "all adult (19 plus years)"
57 55 not 56
58 limit 57 to "all child (0 to 18 years)"
59 55 not 58

Appendix B. Comments on validated outcomes

Pulmonary rehabilitation interventions change the everyday routine of the COPD patients, and generally demand their cooperation and commitment. Similarly, PR interventions require substantial effort from the health care providers and may represent a sizable financial burden. It is therefore very important to assess their efficacy and/or effectiveness using validated outcome measures. Ideally, outcomes ought to be reproducible, sensitive to changes in the measured quantities, to correlate well with the targeted (measured) quantities, and to convey clinically useful information.[1]

We briefly describe and discuss outcome measures for which the minimal clinically significant difference has been estimated. Absolute changes in these outcomes can therefore be perceived as clinically significant or not.

Chronic Respiratory Disease Questionnaire (CRDQ)

The CRDQ has been extensively used in the literature and there is considerable experience from its use.[2] The instrument has four components:

- dyspnoea (assessed in 5 activities the patient deems important). Possible range: 5 to 35
- fatigue (regarding physical function). Possible range: 4 to 28
- emotional function (anxiety, depression). Possible range: 7 to 49
- mastery of breathing (sense of control over disease): Possible range: 4 to 28.

The CRDQ is interviewer-rated and the minimum clinically significant difference is 0.5 points on a 7 point scale.[2] Higher scores are optimal. We have converted CRDQ data on a seven point scale when needed (by dividing the corresponding domains by 5, 4, 7, or 4).

St George's Respiratory disease Questionnaire (SGRQ)

The SGRD instrument is a self-administered 76-item questionnaire (53 questions).[3] The instrument assesses three main areas.

- Symptoms (sputum, coughing, wheezing, dyspnoea)
- Activity
- Impact of disease on daily life

Higher scores imply worst health. The minimum clinically significant difference is 4 units. Dyspnoea is assessed as part of the symptoms domain, but not separately.

6 minute walk test (6MWT)

The 6MWT assesses functional exercise capacity.[4] The patient is instructed to walk for six minutes on zero slope on a standard route, usually back and forth between two signs. The result of the test is the total distance covered in 6 minutes.

It has been shown that the minimal clinically significant difference in the test is approximately 50 meters (54 meters).[4] Differential encouragement may have a great impact on test performance, approximately 30 meters on average (30.4 meters).[5] For this reason it is suggested that standardized encouraging comments should be given by the test supervisor to the patients during testing. It has also been claimed that the 6MWT should be taken three times and the best performance should be recorded.[6] It has also been shown that outdoor and hallway testing may yield different results.[7] The above are described in detail in the ATS guidelines for the 6MWT.[8]

References to Appendix B.

Reference List

(1) Pashkow P, Ades PA, Emery CF, Frid DJ, Houston-Miller N, Peske G et al. Outcome measurement in cardiac and pulmonary rehabilitation. AACVPR Outcomes Committee. American Association of Cardiovascular and Pulmonary Rehabilitation. J Cardiopulm Rehabil 1995; 15(6):394-405.

(2) Guyatt GH, Berman LB, Townsend M, Pugsley SO, Chambers LW. A measure of quality of life for clinical trials in chronic lung disease. Thorax 1987; 42(10):773-778.

(3) Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. Am Rev Respir Dis 1992; 145(6):1321-1327.

(4) Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH. Interpreting small differences in functional status: the Six Minute Walk test in chronic lung disease patients. Am J Respir Crit Care Med 1997; 155(4):1278-1282.

(5) Guyatt GH, Pugsley SO, Sullivan MJ, Thompson PJ, Berman L, Jones NL et al. Effect of encouragement on walking test performance. Thorax 1984; 39(11):818-822.

(6) Steele B. Timed walking tests of exercise capacity in chronic cardiopulmonary illness. J Cardiopulm Rehabil 1996; 16(1):25-33.

(7) Stevens D, Elpern E, Sharma K, Szidon P, Ankin M, Kesten S. Comparison of hallway and treadmill six-minute walk tests. Am J Respir Crit Care Med 1999; 160(5 Pt 1):1540-1543.

(8) ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002; 166(1):111-117.

Appendix C. The Jadad quality scale

The Jadad quality scale[1] was used in the Cochrane review that formed the basis of our reply to question 1. The reader should note that the Jadad scale was not used in the rest of the review.

The Jadad scale is a 3 item/5question scale:

1. Randomization:
    a. Is the study described as randomized?
    b. Is the randomization adequately described/properly generated?
2. Masking:
    a. Is the study described as double blinded?
    b. Is the control treatment described as indistinguishable?
3. Dropouts:
    a. Is there a description of withdrawals?

The total scores range from 0 to 5 points, where trials with 0-2 points are considered to be of poor quality, and those with 3-5 points represent higher quality RCT.[1]

However, as mentioned in the text, pulmonary rehabilitation interventions cannot be double-blinded, and this limits the maximum number of points an RCT can get in the Jadad scale. Thus the applicability of the Jadad scale to RCT in pulmonary rehabilitation interventions is limited. Moreover, it is well appreciated that the use of summary scores from quality scales is problematic. The results are very dependent on the choice of the scale and thus it may be challenging to interpret them. The methodological literature on these topics is extensive.

References to Appendix C.

1) Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Controlled Clin Trials 1996; 17:1-12

Appendix D. Evidence tables for randomized controlled trials not included in the published systematic reviews used in this report

Bauldoff 1996

Country:	US	$N_{centers}$:	1	RCT design:	Parallel
Disease:	COPD			Setting:	Home-based

Interventions, experimental: Arm A

Exercise:	Training (3 upper arm exercises 5/wk; incremental levels) for 8 wk unsupervised, but had supervision once per week.
Other:	No

Interventions, comparator: Arm B

Exercise:	No
Other:	Follow-up by phone calling

Patient flow

N enrolled:	20	Pre-rand exclusions (rationale)	0	N_{rand}	A=	10
					B=	10
Post-rand exclusions (rationale)		0		N_{analyz}	A=	10
					B=	10

Maximum follow-up: End of trial, 8 wk

Population description and baseline data (SD)

Demographics:	Age (y):	A=61(14) B=63(13)	Males (%):	45%	Smokers (%):	ND
Baseline:	FEV_1 (L):	A=0.65(0.37) B=0.96(0.44)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Convenience sample: COPD with FEV_1/FVC <60%, clinically stable,
Exclusion:	Hospitalization within 6 wks, prior arm exercise training, evidence of unstable cardiac disease (MI, HBT, severe CHF), musculoskeletal disability preventing exercise

Quality Assessment for RCT

Blinding:	Assessor	Allocation concealment:	No
Intention-to-treat:	no	Randomization method:	No
Power analysis:	no	Comments:	Overall Quality C

Outcomes at maximum follow-up

Primary:	
Other, efficacy:	Functional exercise capacity test (ring test): Upper extremity endurance test, pre-post data available (SD): Baseline 4 weeks 8 weeks EXP 129 (32) 141 (30) 152 (32) CTRL 139 (46) 138 (48) 141 (43) Significant differences were found for main effects of time (p=0.03), but no significant differences found for main effects of treatment (p=0.95) or interaction of time and treatment (p=0.07)
Other, safety	ND

Berry 1996

Country:	US	$N_{centers}$:	1	RCT design:	3-arms parallel
Disease:	COPD			Setting:	Out-patient

Interventions: Arm A

Exercise:	GER: general exercise reconditioning, consisting of aerobic LLE (walking) and Str for ULE (weight lifting). Intensity 50-75% of maximum Heart Rate, 3/wk for 12 wk
Other:	IMT with threshold device from 15% to 85% of PImax for 15 min two times per day; 7d/wk for 12 wk

Interventions: Arm B

Exercise:	GER (general exercise reconditioning), as in A
Other:	No

Interventions: Arm C

Exercise:	No
Other:	Breathing exercises (pursed lip breathing, and diaphragmatic exercises)

Patient flow

N enrolled:	27	Pre-rand exclusions (rationale)	0	N_{rand}	A= B= C=	9 9 9
Post-rand exclusions (rationale)		2 withdrawals due to acute exacerbations of disease (1 from A, 1 from C)		N_{analyz}	A= B= C=	8 9 8

Maximum follow-up:	6 wk

Population description and baseline data (SD)

Demographics:	Age (y):	A=67.0 (3.4) B=70.8 (6.0) C=70.3 (6.0)	Males (%):	64%	Smokers (%):	ND
Baseline:	FEV_1 [Unit]:	A= not measured But mean should be around 45% to 48% because not different than post data	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	No data					

Patient selection criteria

Inclusion:	FEV_1/FVC <0.65, FEV_1>40% pred; dyspnea on exertion; cigarette or tobacco smoke exposure >20 pack/yrs; ability to self-ambulate, age >60, COPD
Exclusion:	Signif cardiac disease, orthopedic or neurological impairment, serious renal liver or GI disorders, current psych illness or substance dependence, uncontrolled diabetes or HBP, rehab or exercise program w/in 6 mo, saO2 <90% during exercise w/HR >50% age predicted max

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	ND	Allocation concealment:	ND
Intention-to-treat:	No	Randomization method:	ND
Power analysis:	No	Comments:	Overall quality C poor reporting; poor design, very small trial

Outcomes

Primary:	Unclear
Other, efficacy:	Dyspnea: Steady state dyspnea in Borg scale p=0.72 for an ANCOVA among groups. A=2.4 (1.2); B= 2.5 (1.2), C= 2.9 (1.2) Functional exersise capacity (12 min walk distance): There were sign differences among all three groups in an ANCOVA: GER + IMT and for GER compared to control (p=0.03), but no statistically significant differences between GER + IMT and GER.
Other, safety	ND
Comments:	Very small trial, underpowered.

Berry 2003 and Foy 2001

Country:	USA	$N_{centers}$:	Single	RCT design:	Parallel arm
Disease:	Chronic bronchitis and/or emphysema			Setting:	Outpatient

Interventions, experimental; Arm A

Exercise:	Walking and upper-body strength training conducted 3/wk for 3mo
Other:	None

Interventions, comparator (copy-paste if more that 1 arms): Arm B

Exercise:	Walking and upper-body strength training conducted 3/wk for 18 mo
Other:	None

Patient flow

N enrolled:	207	Pre-rand exclusions (rationale)	Total n=67 (failed to complete the exercise orientation; ineligible; dropped out during the run-in period; declined randomization)	N_{rand}	A=	70
					B=	70

Post-rand exclusions (rationale)	In the Short -term (ST) exercise (n=14) In the Long-term (LT) exercise (n=8)	N_{analyz}	A= B=	70 70 (ITT56;62)

Maximum follow-up:	18 mo (First 3 mo both groups underwent supervised, center-based structured exercise therapy; then those randomized into ST no longer continued involvement but were encouraged independent unmonitored home based exercise; the LT group continued exercise training for additional 15 mo)

Population description and baseline data (SD)

Demographics:	Age (y):	A=66.9 B=68.4	Males (%):	A=56% B=56%	Smokers (%):	Current A=34 Current B=33
Baseline:	FEV_1 [L]:	A=1.65 B=1.52	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Other:	FEV_1/FVC (%) 56.4 v 52.3; RV/TLC 55% v 57%					
Comorbidities:	Arthritis (41% v 40%); HTN (41% v 46%); circulatory problems (20% v 13%); heart disease (39% v 34%); diabetes (7% v 4%); any cancer (36% v 24%)					

Patient selection criteria

Inclusion:	Disability with Shortness of breath or diagnosis of chronic bronchitis and/or emphysema; ambulatory; ages 55 to 80 yrs; an expiratory flow limitation FEV1/FVC <70% and FEV1 >20%; not actively engaging in regular exercises or PR for the preceding 6 mo
Exclusion:	Active treatment for cancer; severe CHF; stroke; PVD; CAD; valvuar heart disease; major psychiatric disease; severe anemia; liver or renal disease; uncontrolled diabetes or HTN; orthopedic impairment; blindness or deafness; inability to perform exercise due to physical disability or positive exercise stress test; cognitive impairment; alcohol consumption of >2 drinks/d for the preceding 2 mo.

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	Single (outcome measurement)	Allocation concealment:	ND
Intention-to-treat:	yes	Randomization method:	Stratified blocked randomization (computer generated)
Power analysis:	yes	Comments:	

Outcomes at maximum follow-up

Primary:	Self reported physical disability: 12% less disability in B than those in the A group Adjusted means 1.53 (1.43 to 1.63) v 1.71 (1.61 to 1.81) P=0.016 Physical function B arm patients walked more than 100 feet farther in 6 minutes than those in the A arm +219 ft v +114 ft Δ= 105 ft (P=0.03) B arm patients climbed the two flights of steps 1.3 seconds faster than those in the A arm –0.4 sec v 1.0 sec Δ= -1.4 sec (P=0.05) Time required to move a weight across 6 pegs at shoulder height to be shorter for those in the B arm – 5.4 sec v + 0.5 sec Δ= -5.9 (P=0.06)
Other, efficacy:	Four domains of CRDQ component (mean and SEM) at 18 mo Dyspnea 4.60 (1.43) (short) v 5.10 (0.15) (long) p value for change=0.03 Fatigue 4.53 (1.16) (short) v 5.00 (0.92) (long) p<0.01 Emotional 5.43 (1.04) (short) v 5.56 (0.87) (long) p=0.04 Mastery 6.07 (1.08) (short) v 6.32 (0.85) (long) p=0.04
Other, safety	No deaths occurred during 18 mo
Comments:	

Bourbeau 2003 and 2005 (2 papers)

Country:	Canada	$N_{centers}$:	7	RCT design:	Parallel arm
Disease:	COPD, stable			Setting:	Home-based, unsupervised

Interventions, experimental: Arm A

Exercise:	After the 7th wk, patients were encouraged to exercise at least 3 times weekly for 30-40 min (warm-up and stretching, muscle and cardiovascular exercises [stationary bicycle; walking or climbing stairs]; intensity self regulated at dyspnea of 3 or 4 on the modified Borg scale.
Other:	1 hour weekly teaching at home for 7 or 8 wk on the following topics: teaching about the disease; preventing symptoms; managing an acute exacerbation; adopting healthy lifestyle; leisure activities and traveling.

Interventions, comparator: Arm B

Exercise:	No
Other:	No

Patient flow

N enrolled:	469	Pre-rand exclusions (rationale)	278; 251 refused to participate and 27 lived too far away	N_{rand}	A= B=	96 95
Post-rand exclusions (rationale)		A: 10 dropouts; 5 died, 5 withdrew because of the burden of the evaluation [another 8 died in 2nd year] B: 16 dropouts; 9 died, 1 lost to follow-up and 6 withdrew because of the burden of the evaluation [another 9 died in 2nd year]		N_{analyz}	A= B=	96 95

Maximum follow-up:	24 mo

Population description and baseline data (SD)

Demographics:	Age (y):	A= 69.4 (6.5) B= 69.6 (7.4)	Males (%):	A=50 (52) B= 56 (59)	Smokers (%):	A= 24 (25) B= 25 (26)
Baseline:	FEV_1 [L]:	A=0.98 (0.31) B=1.00 (0.33)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Other:	FEV_1/FVC; drugs taken; education; pack-years of smoking					
Comorbidities:	A vs B [number]: Cardiovascular (41 vs 45); Renal (16 vs 4); Endocrine (18 vs 23); Gastrointestinal (25 vs 30)					

Patient selection criteria

Inclusion:	Patients hospitalized at least once during the previous year for acute exacerbation of COPD; stable COPD for at least 4 weeks before entering study; >50y old; >10 pack-years current or ex-smoker; FEV_1 between 25% and 70% and FEV_1/FVC<0.7.
Exclusion:	Asthma; left congestive heart failure; terminal disease; dementia; uncontrolled psychiatric illness; participation in PR program in the previous year; long-term-care facility stays

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	Blinded test assessor	Allocation concealed:	Yes, central
ITT	Yes	Randomization method	Computer generated
Power analysis:	20% vs 40% difference in hospital admissions with 80% power at 0.05	Comments:	Overall quality A Very well designed, conducted and analyzed

Outcomes at maximum follow-up

Primary:	Hospitalizations for any reason: 92 vs 167 (p<0.01) or –0.7 hospitalizations per pt per year (95% CI –0.95 to –0.46) at the first year; and –0.44 (-0.68 to –0.21) during the 2nd year
Other, efficacy:	• 6MWD did not change significantly neither within nor between groups at 4 or 12 months • SGRQ change: -4.2 (-7.7 to –0.7) at 4 months and -2.0 (-5.9 to 1.8) at 12 months • Mortality: 5 vs 9 at 12 mo follow up; 13 vs 18 at 24 mo follow-up
Other, safety	• Acute exacerbations: at 12 months of follow-up: total 299 vs 366 (p=0.06) • Emergency room visits: -1.3 per pt per yr (-1.18 to -1.42) during the 1st year and -0.7 (-0.58 to -0.82) during the 2nd year • Hospital days per patient: 7.2 (19.5) vs 12.5 (21.2); p=0.01

Brooks 2002

Country:	Canada	$N_{centers}$:	Single	RCT design:	Parallel arm
Disease:	COPD			Setting:	Patients recruited from both in-patient and out-patient; but were given home based routine

Interventions, experimental: Arm A

Exercise:	Enhanced follow-up: invited to attend monthly 2 hr group sessions led by a physical therapist The 1st hour was spent discussing concerns regarding home maintenance program and the 2nd patients performed components of their home program (choice of the patient)
Other:	Additional phone call from a different therapist in between the sessions who asked standardized questionnaire about their adherence to the program

Interventions, comparator: Arm B

Exercise:	Conventional follow-up visited the physical therapist every 3 months for a yr.
Other:	The patients were asked standardized questions regarding their hospitalizations or illnesses. They were asked about their exercise program and encouraged to comply with it. Patients were encouraged to identify their concerns to their therapist and encouraged to resume any of the program that were discontinued

Patient flow

N enrolled:	109	Pre-rand exclusions (rationale)	N=24 (dropped out after baseline evaluation)	N_{rand}	A= B=	48 37
Post-rand exclusions (rationale)		A=13 B=22		N_{analyz}	A= B=	18 23

Maximum follow-up:	12 mo

Population description and baseline data (SEM)

Demographics:	Age (y):	A=68.1±1.1 B=68.1±1.1	Males (%):	A=59 B=58	Smokers (%):	ND
Baseline:	FEV_1 [L]:	A=0.71±0.04 B=0.67±0.04	PaO_2 (mmHg):		$PaCO_2$ (mmHg):	
Other:	6 min walk (m) A=395±15m B=375±14m					
Comorbidities:	ND					

Patient selection criteria

Inclusion:	1) severe stable COPD (forced expiratory volume in one second (FEV_1)<40% predicted, FEV_1/forced vital capacity <0.70); 2) completion of inpatient or outpatient rehabilitation; 3) nonsmoker for a minimum of 6 months; 4) aged 49–85 yrs.
Exclusion:	1) coexisting conditions that might limit exercise tolerance or cognitive functioning; 2) noncompliance with respiratory rehabilitation; 3) mechanical ventilatory support for any part of the day; 4) inability to communicate in English; 5) living too far away to participate.

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	Single (medical follow-up by a respiratory specialist)	Allocation concealment:	ND
Intention-to-treat:	ND	Randomization method:	Stratified (by baseline disability) randomization by random numbers table
Power analysis:	ND	Comments:	

Outcomes at maximum follow-up

Primary:	6 min walk test: Change from baseline A=-53±15m B=-7±5 m using time and group as factors there was no difference in the distance walked in 6 min between the two groups, but a significant difference for time ($p<0.001$) and interaction between time and group ($p=0.03$) Post hoc analysis revealed that for the control group, distances walked at 6, 9 and 12 months were less than the distance at baseline ($p<0.04$). For the EF group, distance walked at 12 months was less than all other measures ($p<0.001$).

	A	CRDQ Dyspnea	4.5±0.2	-0.5±0.1
	B		4.5±0.2	-0.7±0
	A	CRDQ Fatigue	4.7±0.2	-0.7±0.1
	B		4.7±0.2	-0.8±0.1
	A	CRDQ Mastery	5.7±0.2	0
	B		5.4±0.1	0.6±0.1
	A	CRDQ Emotion	5.4±0.2	-0.3±0.1
	B		5.3±0.2	-0.5±0
	There was no difference in total CRDQ score between groups despite a significant difference over time (two-way ANOVA, p=0.32 for group, p<0.001 for time). Post hoc analysis revealed that the quality of life scores at 12 months were lower (worse) than at other times. For the individual domains of the CRDQ, the categories of dyspnoea, fatigue, and mastery showed a difference with time (p≤0.002), but no difference between groups (p >0.1). The category of emotion showed no difference over time or between groups (p >0.1). No significant differences in CRDQ scores were found between groups at 6 months.			
Other, efficacy:	T	SGRQ	42±1	+8±1
	C	Showed effect for time (p=0.002) with no group effect	42±1	+6±1
Other, safety	ND			
Comments:	No			

Chen 1985

Country:	US	$N_{centers}$:	1	RCT design:	Parallel arm
Disease:	COPD			Setting:	Out-patients

Interventions: Arm A

Exercise:	20 minutes cycle ergometer; moderate intensity (unclear) 3/wk for 4wk
Other:	IMT uncontrolled flow, 15 min sessions; intensity up to 62% or PImax; 2/d for 4 wks

Interventions: Arm B

Exercise:	20 minutes cycle ergometer; moderate intensity (unclear) 3/wk for 4wk
Other:	No

Patient flow

N enrolled:	13	Pre-rand exclusions (rationale)	0	N_{rand}	A= B=	7 6
Post-rand exclusions (rationale)		0		N_{analyz}	A= B=	7 6

Maximum follow-up:	End of intervention (4 wk)

Population description and baseline data (SD)

Demographics:	Age (y):	A=62 (11) B=58 (4)	Males (%):	53.8%	Smokers (%):	ND
Baseline:	FEV_1 [%]:	A= 43 (21) B= 40 (20)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	ND					

Patient selection criteria

Inclusion:	ATA criteria, clinically stable
Exclusion:	Other significant disease, exercisability without adverse cardiovascular effects

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	No	Allocation concealment:	No
Intention-to-treat:	No	Randomization method:	No
Power analysis:	No	Comments:	Overall quality C Small trial and poor reporting of methodological quality items

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	Figure data, converted to Watts and min, after digitizing and making calculations, difference in the mean change from baseline: ICET Wmax: CWR at 66% of Wmax A-B= 0.8 W (-23.2, to 24,8) 0.8 min (-4.0, 2.3)
Other, safety	ND

Dekhuizen 1991

Country:	Holland	$N_{centers}$:	1	RCT design:	Parallel
Disease:	COPD			Setting:	Out-patient

Interventions, experimental: Arm A

Exercise:	Cycling, walking, back, shoulder and abdominal muscle exercises (endurance training) for 2h; intensity at 80% of maximum heart rate; 5/wk for 10 wk
Other:	Breathing retraining; Education; Relaxation techniques + IMT with controlled flow for 15 min twice a day; intensity at 70% PImax, reset twice per week; per day for 10 wk

Interventions, comparator: Arm B

Exercise:	Cycling, walking, back, shoulder and abdominal muscle exercises (endurance training) for 2h; intensity at 80% of maximum heart rate; 5/wk for 10 wk
Other:	Calisthenics (energy conservation and work minimization?); Breathing retraining ; Education; Relaxation techniques

Patient flow

N enrolled:	40	Pre-rand exclusions (rationale)	0	N_{rand}	A= B=	20 20
Post-rand exclusions (rationale)		0		N_{analyz}	A= B=	20 20

Maximum follow-up:	End of intervention (10wk)

Population description and baseline data (SD)

Demographics:	Age (y):	A=60 (7) B=58 (8)	Males (%):	75%	Smokers (%):	ND
Baseline:	FEV_1 [%]	A=51.7 (17.0) B=46.9 (14.0)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	A=38.5 (3.9) B=38.5 (3.9)
Other:	FRC%TC, TLD pred					
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Moderate to severe airflow obstruction; normal resting arterial PCO_2 that increased during exercise; patients whose alveolar-arterial O_2 presure difference does not increase more than 15 mmHg after maximal progressive exercise (comment: excludes people with severe V-Q mismatching, people with shunts or diffusion phenomena during exercise.)
Exclusion:	ND

Quality Assessment for RCTs

Blinding:	NS	Allocation concealment:	NS
Intention-to-treat:	NS	Randomization method:	NS
Power analysis:	NS	Comments:	Overall quality B small trial; Selected patient population, not ANCOVA or MANOVA

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	Functional exercise capacity: 12 min walking distance improved in both arms but significantly more in IMT ($p<0.05$ in Mann-Whitney comparing the post-treatment results) However, difference in change from baseline: is 69m (-141, 279) Maximal exercise capacity: Maximal work load on bicycle ergometer increased significantly in both arms ($p<0.01$) Difference in change from baseline in maximum work (or *power*, W) is -7W (-14.4, 0.4)
Other, safety	ND

Elliott 2004

Country:	Australia	$N_{centers}$:	2	RCT design:	2 short term or 3-armed long term parallel
Disease:	COPD			Setting:	various

Comment: They have employed a peculiar design that has 2 phases:
First a 3 month phase and then another, long term maintenance program.
In the end the arms were (short;3mo/long;upto 12mo)
First Arm: Hospital/hospital
Second Arm: Hospital/community
Third arm: Community/community
Because of the high dropouts (16/43 completed the 12 months) the authors have not analyzed long-term results.
We present short term results of hospital vs community care, merging the first two arms for the short term outcomes

Interventions, experimental; Arm A

Exercise:	Hospital-based program, ULE + LLE (END) and torso strengthening for 1.5 h (continually supervised by physiotherapist); intensity unclear; 2/wk for 3 months
Other:	

Interventions, comparator Arm B

Exercise:	Community-based program with general exercises (aerobic) for 1.5 h; intensity unclear; 2/wk fot 3 months
Other:	

Patient flow

N enrolled:	43	Pre-rand exclusions (rationale)		N_{rand}	A= B=	30 13
Post-rand exclusions (rationale)		12 dropouts (5 illness, 4 lack of interest, 1 transport, 1 preference for other regimen, 1 pt was "too well".		N_{analyz}	A= B=	22 9

Maximum follow-up:	End of intervention

Population description and baseline data

Demographics:	Age (y):	A=67.5 average B=62.5 average	Males (%):	53.5	Current Smokers (%):	A=93.3 B= 77
Baseline:	FEV_1 [%]:	A=46.2 average B=42.7 average	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	3 pts w/asthma, 1 w/bronchiectasis					

Patient selection criteria

Inclusion:	Moderate to severe COPD (FEV1 34-70% of predicted)
Exclusion:	Cardiac or other disease, musculoskeletal problems, significant arterial O2 desaturation during exercise, difficulty with communication or recent respiratory infections

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	NS	Allocation concealment:	ND
Intention-to-treat:	No, but "results were the same"	Randomization method:	ND
Power analysis:	NS	Comments:	Overall Quality C; Peculiar design; difficult to understand what key question they targeted with this design

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	• CRDQ (total score): No significant difference between the 2 groups (ANOVA), but SS improvement in both. For A vs B, difference in changes from baseline was 6.1 (-5 to 17) • Functional exercise capacity: 6MWT (m): SS in hospital based only; Difference in changes from baseline was approximately 70m (16 to 123)
Other, safety	ND

Emery 1998

Country:	USA	$N_{centers}$:	1	RCT design:	Parallel arm
Disease:	COPD, stable			Setting:	Out-patient

Interventions: Arm A

Exercise:	Lower limb exercise (endurance: stationary cycling, walking) + upper limb exercise for 45 min; intensity unclear; daily for 5wk; out-patient; then 3/wk for 60 to 90 min for 5wk
Other:	Stress management (1h/wk) and education (4h per week)

Interventions: Arm B

Exercise:	No
Other:	Stress management (1h/wk) and education (4h per week)

Interventions: Arm C

Exercise:	No
Other:	No

Patient flow

N enrolled:	92	Pre-rand exclusions (rationale)	13 (9 with normal FEV1, 4 could not commit to the program)	N_{rand}	A= 29 B= 25 C= 25
Post-rand exclusions (rationale)		A=4 dropped because of illness B=2 dropped because of transportation problems C=0		N_{analyz}	A= 25 B= 23 C= 25

Maximum follow-up:	End of trial, 10wk

Population description and baseline data (SD)

Demographics:	Age (y):	A=65 (6) B=67 (6) C=67 (7)	Males (%):	A=15 (52) B=10 (71) C=12 (80)	Smokers (%):	ND
Baseline:	FEV_1 [L]:	A=1.24 (0.6) B=1.13 (0.5) C=1.02 (0.4)	PaO_2 (mmHg):	A=75.4 (12.7) B=76.0 (9.5) C=72.5 (8.3)	$PaCO_2$ (mmHg):	ND
Other:						
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Age >50y, FEV_1/FVC<.7, clinical symptoms of COPD for at least 6 months.
Exclusion:	Asthma, TB, pulmonary fibrosis, cancer, cardiac disease, medical conditions limiting ability to participate in an exercise program.

Quality Assessment for RCT

Blinding:	Baseline assessment	Allocation concealment:	Unclear
Intention-to-treat:	No	Randomization method:	Random number tables
Power analysis:	No	Comments:	Overall Quality B

Outcomes at maximum follow-up

Primary:	unclear
Other, efficacy:	The work in kP*m (pre to post): A=66.3 (29.8) to 77.6 (34.8) (p<0.01 for pre-post) B=64.1 (23.6) to 65.9 (22.9) C=59.2 (24.9) to 59.1 (27.7) The difference in the mean change from baseline would be: A vs B= 9.5 (-7.0, 26.0) - 93 J (-69, 255) A vs C=11.4 (-6.0, 28.8) - 112 J (-59, 282)
Other, safety	ND

Finnerty 2001

Country:	UK	$N_{centers}$:	1	RCT design:	Parallel Arm
Disease:	COPD, stable			Setting:	Home based/Outpatient

Interventions, experimental: Arm A

Exercise:	Supervised aerobic training 1h/week for 6 wks (UL, LL). Patients were also asked to exercise (unsupervised) daily for 5 days/wk using a walking program with 9 levels (maximum was 10 min of walking, 10 min of rest)
Other:	Discussion of breathing techniques; educational program, 2h/week for 6 wks; weekly dietary advice; 2 occupational therapy sessions on coping with loss of interest in leisure activities due to breathlessness; discussions on coping with anxiety, sleep problems and on relaxation techniques

Interventions, comparator: Arm B

Exercise:	No
Other:	3 visits to outpatient facilities, one every month

Patient flow

N enrolled:	100	Pre-rand exclusions (rationale)	0	N_{rand}	A= 50 B= 50	
Post-rand exclusions (rationale)		(A) 10 pts failed to attend assessment; 4 withdrawn (non-COPD); 4 did not complete; additional 8 not assessed at 24 wk (B) 17 pts failed to attend assessment; 4 withdrawn (non-COPD); 4 did not complete		N_{analyz}	A= 32 B= 23	

Maximum follow-up:	24 wks

Population description

Demographics:	Age (y):	A=70.4 (8.0) B=68.4 (10.4)	Males (%):	44 (67), overall, starting	Smokers (%):	8 (12)
Baseline:	FEV_1 [Unit]:	A= 41.2% (19.2) B=41.2% (16.2)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Stable COPD patients who had their therapy optimized
Exclusion:	Dementia, marked agitation or depression evident to the investigators, congestive heart failure, cor pulmonale, malignancy, cerebrovascular accident

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	Blinded assessor for SGRQ, 6MWT	Allocation concealed	NS
ITT	NS	Randomization method	NS
Power analysis:	NS	Comments:	Overall quality C High attrition rates

Outcomes

Primary:	SGRQ (total score), change from baseline favoring PR= -8.1 (-14.9 to –1.5) at 24 wks Statistically significant changes from baseline in all three SGRQ subscores (symptom, activity, impact; $p<0.01$ for all three)
Other, efficacy:	6MWT: Change from baseline= 51m (20 to 81) at 12 wks, and 53m ($p>0.05$) at 24 wks
Other, safety	1 pnt in the Comp arm missed the 3rd session “because of illness”

Foglio 2001

Country:	Italy	$N_{centers}$:	Single	RCT design:	Parallel arms
Disease:	COPD and chronic bronchial asthma			Setting:	Outpatient

Interventions, experimental: Arm A

Exercise:	30 patients underwent another a second PRP (PRP2) and related post-PRP evaluations 1-year after PRP-1
Other:	At the end of the second year, patients underwent clinical and physiologic evaluations and underwent the third PRP (PRP3)

Interventions, comparator: Arm B

Exercise:	31 patients (control) did not undergo PRP2
Other:	At the end of the second year, patients underwent clinical and physiologic evaluations and underwent the third PRP (PRP3)

Patient flow

N enrolled:	61	Pre-rand exclusions (rationale)	None	N_{rand}	A= B=	30 31
Post-rand exclusions (rationale)		All completed PRP2 but did not enter PRP3 (N=25). Eleven patients in group A and 10 patients in group B did not perform evaluations at T4 due to personal, transport, or familial problems. Four more patients (two patients in each group) were excluded from the study due to intervening pathologic conditions (one bladder cancer, two limb traumas, one sudden onset of ischemic heart disease).		N_{analyz}	A= B=	17 19 (yr 2)

Maximum follow-up:	2 yrs

Population description and baseline data (SD)

Demographics:	Age (y):	A=61±8 B=59±9	Males (%):	A=71% B=42%	Smokers (%):	ND
Baseline:	FEV_1 [Unit]:	A=64±19 B=69±31	PaO_2 (mmHg):	A=78±13 B=77±8	$PaCO_2$ (mmHg):	A=42±14 B=39±2
Other:	BDI A=7.5±1.3 B=8.3±1.5		SGRQ A=38±16 B=33±20		6 MWD m A=439±114 B=485±68	
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Consecutive and stable chronic airway obstruction
Exclusion:	Patients with organ failure, or cancer or who were unable to cooperate

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	Single blind (data collector)	Allocation concealment:	ND
Intention-to-treat:	ND	Randomization method:	ND
Power analysis:	ND	Comments:	No

Outcomes at maximum follow-up

Primary:	Exercise tolerance measured by peak work rate and 6 MWD: patients of both groups showed an increase in exercise tolerance as assessed by peak work rate and 6MWD, which was lost at T2. At T3, in group A exercise tolerance increased again. At T4, this benefit was lost again, such that exercise tolerance in patients of group A was not significantly different from control subjects who did not attend any other PRP but PRP1. PRP3 resulted in a new improvement in exercise tolerance for both groups (Shown in fig)
Other, efficacy:	Each PRP was followed by an improvement in the TDI, but no difference was observed between the two groups at any time in either index. Each PRP was followed by a significant short-term improvement in dyspnea at isoworkload, as assessed by the Borg scale; but at T4, no difference was observed between the two groups and with T0. (Values for change not correctly reported in table 2)
Other, safety	In the second-year follow-up, the substantial lack of hospitalizations observed in the first year following PRP1 was maintained in both groups, independent of the participation to PRP2. In the year following PRP2, 8 of 17 patients of group A but 0 of 19 control patients suffered from no exacerbation. The difference was significant. This means also significant further reduction in exacerbations observed in group A but not in group B in comparison to the first year after PRP1.
Comments:	Most of the results shown in figures

Goldstein 1989

Country:	Canada	$N_{centers}$:	1	RCT design:	Parallel arm
Disease:	Stable COPD			Setting:	In-patient

Interventions, experimental: Arm A

Exercise:	8-wk inpatient rehabilitation program with treadmill walking and interval training for upper and lower extremities; unclear intensity and frequency
Other:	Education, breathing retraining, relaxation classes + sham IMT

Interventions, comparator: Arm B

Exercise:	8-wk inpatient rehabilitation program with treadmill walking and interval training for upper and lower extremities; unclear intensity and frequency
Other:	Education, breathing retraining, relaxation classes + IMT with pressure threshold device; begin with load sustainable for 10 minutes and then train till it is sustainable for 20 min, then repeat cycle; twice per day, 5d/wk for 4wk

Patient flow

N enrolled:	12	Pre-rand exclusions (rationale)	0	N_{rand}	A= B=	6 6
Post-rand exclusions (rationale)		One from A group discontinued because of infection		N_{analyz}	A= B=	5 6

Maximum follow-up:	End of intervention (4wk)

Population description and baseline data (SD or SEM)

Demographics:	Age (y):	A= 65 (7) B= 66 (7)	Males (%):	10 (83)	Smokers (%):	ND
Baseline:	FEV_1 [%]:	A=27 (10) B=38 (13)	PaO_2 (mmHg):	A=62 (7) B=70 (11)	$PaCO_2$ (mmHg):	A=42 (7) B=37 (7)
Other:	VC, FEV1/FVC, TLC, Dsb					
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Patients with severe but stable COPD participating in a PR program and highly motivated.
Exclusion:	Unclear

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	NS	Allocation concealment:	NS
Intention-to-treat:	No	Randomization method:	NS
Power analysis:	No	Comments:	Overall quality C too small study sample size, sham IMT identified by patients

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	Functional exercise capacity, 6MWT: Difference in the change from baseline 64m (-70, 198) Endurance - symptom limited response at 10 min at baseline. Difference in the change from baseline 1.1 min (-4.9, 7.1)
Other, safety	ND
Comments:	This is earlier-published compared to Goldstein's other PR paper, so overlap is unlikely.

Larson 1999

Country:	US	$N_{centers}$:	1	RCT design:	Single blind, 4-armed Parallel
Disease:	COPD			Setting:	Home-based

Interventions, Arm A

Exercise:	IMT training using threshold loaded device for 30 min per day. Interval training protocol; starting at 30% till 60% of PImax; 5d / week for 4 months.
Other:	No

Interventions, Arm B

Exercise:	CET: Interval training protocol. For 20 min per day; Begin at 50% of maximum work rate and go on till 85% of max predicted HR or till symptoms limit exercise; 5d /wk, for 4 months. Nurse weekly home visits ensure compliance.
Other:	No

Interventions, Arm C

Exercise:	IMT + CET, as above
Other:	No

Interventions, Arm D

Exercise:	ED, 8 home visits every other week.
Other:	No

Patient flow

N enrolled:	130	Pre-rand exclusions (rationale)	Unclear which of the below quoted are pre rand. Seems that most are post rand!	N_{rand}	A= B= C= D=	? ? ? ?
Post-rand exclusions (rationale)		Exacerbation of lung disease 8, other health problems 9, no interest 6, inability to exercise 4, poor adherence 3, response to graded exercise test 34, other 13. Those who were disqualified as result of graded exercise test: CV problems 23, cd not tolerate testing 7, had oxyhemoglobin desaturation 2, orthopedic problems 2.		N_{analyz}	A= B= C= D=	13 14 14 12

Maximum follow-up:	4 mo

Population description and baseline data (SD)

Demographics:	Age (y):	A=66 (5) B=66 (6) C=68 (6) D=62 (7)	Males (%):	66%	Smokers (%):	ND
Baseline:	FEV_1 % pred	A=55 (17) B=46 (17) C=46 (17) D=55 (18)	PaO_2 (mmHg):	A=77 (8) B=76 (6) C=74 (8) D=78 (9)	$PaCO_2$ (mmHg):	A=40 (4) B=39 (3) C=41 (3) D=39 (4)
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Age 45-75, moderate to severe air flow obstruction (FEV <65% pred and FEV/FVC <70%, complaints of dyspnea on exertion, clinically stable, no pulmonary rehab w/in 1 yr
Exclusion:	Evidence of asthma, major exacerbation 2 mo before enrollment, >10 mg/d prednisone, home oxygen therapy, oxyhemoglobin saturation <85% w/exercise, other health problems interfering with exercise

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	single	Allocation concealment:	ND
Intention-to-treat:	no	Randomization method:	ND
Power analysis:	no	Comments:	Overall Quality C Too high attrition rate Very small trial with 4(!) arms

Outcomes at maximum follow-up

Primary:	Maximal exercise capacity, ICET (calculated from pre-post measurements): • CET + IMT vs IMT: 15W (-4, 34) • CET + IMT vs CET: 0W (-23, 23) • CET vs ED: 15W (-7, 37) • CET vs IMT: 15W (-4, 34)

	• CET and CET+IMT vs IMT and EDU p<0.05 (ANOVA?)
Other, efficacy:	CRQ dyspnea (converted to 7 point scale, calculated from pre-post measurements): • CET + IMT vs IMT: -0.1 (-1, 0.8) • CET + IMT vs CET: -0.3 (-1, 0.4) • CET vs ED: 0.7 (0, 1.4) (p>0.05) • CET vs IMT: 0.2 (-0.7, 1) • MANOVA group effect not significant (for the 2 CRQ scores) CRQ fatigue (converted to 7 point scale, calculated from pre-post measurements): • CET + IMT vs IMT: 0.5 (-0.3, 1.1) • CET + IMT vs CET: 0.2 (-0.6, 0.9) • CET vs ED: 0.2 (-0.7, 1.1) • CET vs IMT: 0.4 (-0.5, 1.2) • MANOVA group effect not significant (for the 2 CRQ scores) Functional exercise capacity, CET (calculated from pre-post measurements): • CET + IMT vs IMT: -2W (-19, 15) • CET + IMT vs CET: 3W (-16, 22) • CET vs ED: -3W (-23, 17) • CET vs IMT: -5W (-21, 11) • CET and CET+IMT vs IMT and EDU p<0.01 (ANOVA?)
Other, safety	ND

Mador 2005

Country:	US	$N_{centers}$:	1	RCT design:	Parallel
Disease:	Stable COPD			Setting:	In-patient

Interventions: Arm A

Exercise:	Endurance training in the cycle ergometer; intensity at 50% of W max in the incremental cyclea ergometry test. If patients could sustain 20 min with Borg dyspnea <5, workload was increased (by 10%), and this was for 8wk.
Other:	Education 1h/wk

Interventions: Arm B

Exercise:	Endurance training like in arm A
Other:	Education 1h / wk + IMT with hyperpnea training (rebreathing bag) with visual feedback; supervised; frequency unclear

Patient flow

N enrolled:	38	Pre-rand exclusions (rationale)		N_{rand}	A= B=	19 19
Post-rand exclusions (rationale)		5 patients in A and 3 in B failed to complete the rehabilitation program, 1 patient in B refused post-rehabilitation measurement		N_{analyz}	A= B=	14 15

Maximum follow-up:	End of follow-up (8wk)

Population description and baseline data (SD)

Demographics:	Age (y):	A=70.9 (7.5) B=69.7 (7.7)	Males (%):	ND	Smokers (%):	Current 0; Ex probably all
Baseline:	FEV_1 [%]	A=44 (15.5) B=45 (22.4)	PaO_2 (mmHg):	A=79.9 (11.6) B=75.4 (10.5)	$PaCO_2$ (mmHg):	A=39.1 (8.2) B=37.5 (5.0)
Comorbidities:	ND					

Patient selection criteria

Inclusion:	COPD diagnosed clinically, successfully quit smoking within 3 mo.
Exclusion:	Unclear

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	yes	Allocation concealment:	ND
Intention-to-treat:	no	Randomization method:	Cluster randomized (in classes of 3-5); Actual method ND
Power analysis:	no	Comments:	Overall Quality C Failure to account for intracluster correlation; unclear cluster size and cluster formation procedure

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	CRDQ domains difference in the mean change from baseline (B vs A): Dyspnea: -0.64 (-1.32, 0.04) Fatigue: -0.65 (-1.35, 0.05) Emotional: -0.16 (-0.89, 0.57) Mastery: -0.42 (-1.08, 0.24) Both arms had statistically significant improvement in all four domains ($p<0.03$), except for the emotion and mastery domains in group B ($p>0.05$). Functional exercise capacity: 6MWT: B vs A difference in the mean change from baseline is 4m (-70, 78) Both groups improved statistically significantly compared to baseline (but on average only 40 and 44 m in A and B respectively) Endurance at 60-70% of Wmax in ICET: B vs A difference in the mean change from baseline is –2.4 min (-12.0 to 7.2) Endurance exercise time increased significantly in both arms ($p<0.005$ for both arms).

	Maximal exercise capacity, ICET: B vs A difference in the mean change from baseline is 0W (-15 to 15) Only group A improved significantly from baseline (p=0.0036); for group B the corresponding p-value was 0.07.
Other, safety	ND

McKeon 1896

Country:	Australia	$N_{centers}$:	1	RCT design:	Parallel arm
Disease:	COPD			Setting:	Home-based

Interventions: Arm A

Exercise:	General exercise training (graded walking for 12 min, and then stair climbing); intensity unclear; each day for 6 weeks.
Other:	Education and breathing exercises (no details) + IMT training, breathing through resistance for 15 minutes per day, uncontrolled flow.

Interventions: Arm B

Exercise:	General exercise training (graded walking for 12 min, and then stair climbing); intensity unclear; each day for 6 weeks.
Other:	Education and breathing exercises (no details) + sham IMT training

Patient flow

N enrolled:	18	Pre-rand exclusions (rationale)	0	N_{rand}	A= B=	10 8
Post-rand exclusions (rationale)		0		N_{analyz}	A= B=	10 8

Maximum follow-up: End of trial (6 wk)

Population description and baseline data (SD)

Demographics:	Age (y):	A= 67 (7.8) B= 69 (8.7)	Males (%):	ND	Smokers (%):	ND
Baseline:	FEV_1 [%]:	A=33 (9) B=39 (8)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Severe airflow obstruction
Exclusion:	Cardiac disease or other significant medical conditions

Quality Assessment for RCTs

Blinding:	No	Allocation concealment:	ND
Intention-to-treat:	NA (no dropouts)	Randomization method:	ND
Power analysis:	No	Comments:	Overall Quality C Very poor reporting of methodological quality items, small trial

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	Functional exercise capacity, 12MWT: There were no statistically significant changes from baseline in either arm. There were no statistically significant differences between arms (arm A from 733 to 711m on average before and after training, and arm B from 819 to 856m of average, before and after treatment, respectively). Endurance, stair climbing: There was a clinically unimportant but statistically significant increase in the number of stairs climbed until exhaustion for arm A (from 64 to 70 stairs on average before and after training [$p<0.05$]). A non-significant tendency was also observed for arm B (from 67 to 73 stairs on average, before and after treatment, respectively). There were no statistically significant differences between arms for this outcome.
Other, safety	ND

Newall 2005

Country:	UK	$N_{centers}$:	1	RCT design:	3-armed parallel
Disease:	Idiopathic bronchiectasis			Setting:	Outpatient

Interventions: Arm A

Exercise:	Exercise training 3/wk at 80% of peak HR, 2 at hospital, 1 at home
Other:	IMT with pressure threshold device at individually programmed intensity, 15 min twice per day

Interventions: Arm B

Exercise:	Exercise training 3/wk at 80% of peak HR, 2 at hospital, 1 at home
Other:	Sham IMT training

Interventions: Arm C

Exercise:	No intervention
Other:	No intervention

Patient flow

N enrolled:	32	Pre-rand exclusions (rationale)		N_{rand}	A= B= C=	11 12 9
Post-rand exclusions (rationale)		During training, 1 withdrew from A and 1 from B due to exacerbation of disease; during follow-up, 4 withdrew from A and B (2 for personal reasons, 2 due to exacerbation of disease)		N_{analyz}	A= B= C=	10 7 9

Maximum follow-up:	8 wk trial, 3 mo FU

Population description and baseline data (SD)

Demographics:	Age (y):	A=57.3 (2.4) B=63.1 (3.5) C=62.9 (3.9)	Males (%):	23.1	Ex-Smokers (%):	A=20 B=57.1 C=22.2
Baseline:	FEV_1 [Unit]: L	A=1.23 (0.74) B=1.44 (0.77) C=1.49 (0.61)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Idiopathic bronchiectasis confirmed by hi-resolution computed tomography
Exclusion:	Evidence of concomitant emphysema in high resolution CT, endocrine, orthopedic or primary cardiac disorders, CAD, hypertension, cor pulmonale. Any acute exacerbation within previous 6 wk or undergoing long-term oral corticosteroids.

Quality Assessment for RCTs

Blinding:	No	Allocation concealed	Yes, central randomization
ITT	No	Randomization method	Computer generated
Power analysis	Yes, 11 per arm for 80% power at alpha=0.05 to detect 50m difference in ISWT (SD=40)	Comments:	Overall quality B Small sample sizes; adequately powered to find a difference

Outcomes at maximum follow-up

Primary:	PImax and ISWT; For ISWT distance (difference in change from baseline): • A vs B: at the end of training 27.8m (-43.9 to 99.5), and at 3 months of follow up: 75.7m (-9.7 to 161.0) (figure 3 digitized data) • A vs C: at the end of training 113.5m (46.2 to 180.8)
Other, efficacy:	• Endurance treadmill test at 85% of peak O_2 uptake: difference in change from baseline (m) A vs B: at the end of training 214.5m (-7.4 to 436.4), and at 3 months of follow-up: 515m (12 to 1020) favoring A (Figure 2 digitized data) A vs C: at the end of the training 720m (279 to 1161) favoring A • Total score in SGRQ: A vs B: At 3 mo, the difference in the mean change from baseline was -12.3 (-24.7, 0.1) A vs C: At the end of training arm A had mean change from baseline -7.7 (95% CI: -16.6 to 1.1; reported as better that that of C with p=0.05). After 3 mo A had a mean change from baseline -10.0 (95% CI: -21.3 to 1.3; reported as "statistically significant" [compared to C?])
Other, safety	ND

Norweg 2005

Country:	US	$N_{centers}$:	1	RCT design:	3-arm parallel trial
Disease:	COPD			Setting:	Out-patient + home-based

Interventions: Arm A

Exercise:	exercise training (LLE+ ULE, endurance) 15 1-hr sessions; 2/wk for 3wk; and encouraged to exercise at home for 20 to30 min, 2 or 3 times per week
Other:	

Interventions: Arm B

Exercise:	B=exercise training
Other:	activity training: structured behavioral intervention emphasizing dyspnea management strategies especially controlled breathing combined with supervised activity exertion, 6 1 hr/wk sessions

Interventions: Arm C

Exercise:	exercise training as in A
Other:	Lecture series

Patient flow

N enrolled:	67	Pre-rand exclusions (rationale)	0	N_{rand}	A= B= C=	18 10 15
Post-rand exclusions (rationale)		There are 6 wk, 12 wk, 18 wk, and 24 wk evaluations Attrition was due to COPD related surgery, illness, injury, finding the program intensity too great, being unreachable or unwilling to cooperate.		6 wk A=11 B=10 C=12	12 wk A=9 B=10 C=11	24 wk A=6 B=8 C=7

Maximum follow-up:	24 wk

Population description and baseline data (SD)

Demographics:	Age (y):	A=77.1(4.0) B=73.5(4.5) C=70.1(7.3)	Males (%):	51.2	Smokers (%): Max pks/d	A=1.1 B=1.5 C=1.9
Baseline:	FEV_1 [%]:	No baseline data, but was 55% at 6 wk	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Medically stable outpatients with COPD, aged >60
Exclusion:	Cognitive deficits (MMSE score < 24), dementia, blindness, unstable angina, other disabling condition

Quality Assessment for RCT

Blinding:	ND	Allocation concealment:	ND
Intention-to-treat:	no	Randomization method:	Minimization: ("biased coined design and probability tables")
Power analysis:	no	Comments:	Overall quality B Analyzed with mixed models, but nevertheless, sample too small. Imbalance in age and CRDQ domains between groups; high attrition rates, very small study

Outcomes

Primary:	Unclear
Other, efficacy:	CRDQ, adjusted mean differences across groups (adjusted for age and baseline values)

	B vs A		C vs A	
	difference	p-value	difference	p-value
Dyspnea	0.69	>0.05	-0.28	>0.05
Fatigue	0.63	>0.05	0.11	>0.05
Emotion	-0.12	>0.05	-0.82	<0.05
Mastery	-0.28	>0.05	-0.68	>0.05

No significant interaction effects of time were found for CRDQ domains (p>0.27). Mean emotional function scores of B (p=0.02) and A (p=0.03) were signif better than C. No main treatment effects were found for dyspnea (p=0.09), fatigue (p=0.22), mastery (p=0.37)
No significant differences found for 6 min walk distance (p=0.77).
Adjusted post-training walking distances improved by 55m (p<0.0001, adjusted for treatment and time).

Other, safety	ND

Porta 2006

Country:	Italy	$N_{centers}$:	3	RCT design:	Parallel arm
Disease:	Patients recently weaned from mechanical ventilation			Setting:	Inpatient (RICU)

Interventions, experimental: Arm A

Exercise:	ULE (cycling, arm ergometer) for 20min/d for 15 sessions; intensity titrated on dyspnea scores
Other:	General physiotherapy: 45 min/session, six sessions /wk; components included deambulation, trunk efficiency and cough control, chest physiotherapy, functional and strengthening exercises (considered usual care in RICU).

Interventions, comparator: Arm B

Exercise:	No
Other:	As in the intervention arm

Patient flow

N enrolled:	66	Pre-rand exclusions (rationale)	0	N_{rand}	A= B=	32 34
Post-rand exclusions (rationale)		Arm A: 7 dropouts. 1 nosocomial pneumonia; 3 Acute Respiratory Failure (ARF) without infection; 3 joint/muscle pain Arm B: 9 dropouts. 3 nosocomial pneumonia; 4 ARF without infection; 1 joint/muscle pain; 1 abdominal pain		N_{analyz}	A= B=	32 34

Maximum follow-up:	unclear

Population description and baseline data (SD)

Demographics:	Age (y):	A=70 (5.6) B=72 (5.2)	Males (%):	A=22 (69) B=23 (68)	Smokers (%):	ND
Baseline:	FEV_1 [Unit]:	A= 44% (20) B= 43% (24)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	A=50 (10.5) B=54 (10.7)
Other:	ND					
Patient diagnoses:	83% tracheostomy; N= 46 COPD, 10 restrictive chest wall disease (6 fibrothorax; 4 kyphoscoliosis); 6 cardiosurgical sequelae; 2 sepsis; 1 thoracic trauma; 1 abdominal surgery; 31 were on LTOT					

Patient selection criteria

Inclusion:	Patients weaned from mechanical ventilation (MV) after >48 to <96 h; clinically stable by the values of arterial blood gases; no fever or infection; stable hemodynamics; conscious and cooperative mental state. Definition of weaning: >48 hours of spontaneous breathing without the following: respiratory rate > 35 breaths/min; PaO_2<50mmHg at FiO_2 >=40%; heart rate >135; major arrhythmias requiring IV Rx; SBP >180 mmHg or <70 mmHg; agitation and anxiety; new appearance of diaphoresis
Exclusion:	Primary neurologic disease; CVD; myopathy, cardiovascular instability, severe arrhythmia, orthopedic problems, insufficient cooperative state

Quality Assessment for RCTs

Blinding:	NS	Allocation concealment:	NS
Intention-to-treat:	Yes	Randomization method:	NS
Power analysis:	No	Comments:	Overall quality B

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	Difference in change from baseline: • Maximal exercise capacity: 4.7W (1.69, 7.75) • Borg dyspnea after maximal exercise test: 0.26 (-0.97 to 1.49) • Functional exercise capacity, at 50% of peak work rate: 4.12 min (0.68, 7.56) • Borg dyspnea after functional exercise test: -1.46 (-2.93, 0.014)
Other, safety	No

Puente-Maestu 2000 (2 papers)

Country:	Spain	$N_{centers}$:	1	RCT design:	Parallel
Disease:	COPD			Setting:	In-patient and home-based (see arms)

Interventions: Arm A

Exercise:	Home-based exercise, self-monitored with a pedometer; intensity 3-4 km in 1 h; 4d/wk for 8 wk
Other:	No

Interventions: Arm B

Exercise:	Physiotherapist-supervised training on a treadmill for 60 min; at 80% of highest O_2 consumption without lactic acidosis emerging, or 50% if the lactic acidosis threshold was not found; 4d/wk for 8 wk
Other:	No

Patient flow

N enrolled:	49	Pre-rand exclusions (rationale)	No	N_{rand}	A= B=	24 25
Post-rand exclusions (rationale)		Arm A=3 dropouts and Arm B =5 dropouts, all for scheduling reasons or personal affairs. Additional 4 patients from A and 2 from B did not produce breath by breath signals of quality to study the kinetics.		N_{analyz}	A= B=	17 18

Maximum follow-up:	End of intervention, 8wk

Population description and baseline data

Demographics:	Age (y):	A=63.4 (4.8) B=65.8 (5.7)	Males (%):	100	Smokers (%):	0
Baseline:	FEV_1 [Unit]:L	A=1.09 (0.17) B=1.09 (0.19)	PaO_2 (mmHg):	A=67.5 (5.4) B=62.8 (8.5)	$PaCO_2$ (mmHg):	A=37.9 (2.6) B=37.7 (3.3)
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Nonsmoking males <75 yr w/severe COPD, history of smoking 10 pks/yr; declared smoking cessation w/in 6 mo; stable phase of COPOD meaning no exacerbation at least 2 mo or acute dyspnea needing med assistance or changes in volume of sputum, increase in lung sound; >grade 2 dyspnea; post-bronchodilator FEV <50% predicted value; <15% increase in FEV after bronchodilation; <3% carboxyhemoglobin;
Exclusion:	Other significant lung or extrapulmonary dis, or physical disability; asthma, bronchiectasis; obliterating bronchiolitis, scarring affecting >20% hemithorax in chest radiography, thoracic deformity, fibrothorax, severe cardiomyopathy, ischemic cardiopathy, severe arryhthmia, type I diabetes, neuromuscular disorders, severe hepatic or renal diseases

Quality Assessment for RCTs

Blinding:	Assessor	Allocation concealment:	NS
Intention-to-treat:	No	Randomization method:	Unclear method, blocks of 4
Power analysis:	NS	Comments:	Overall quality C Inconsistencies between 2 papers, 30% attrition rates

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	• CRDQ (n=41) No differences before or after training. Scores in all 4 dimensions improved significantly. No significant differences in magnitude of change or proportion of patients who improved the score by any clinically significant amount. • Functional exercise capacity, (n=35) Constant work-rate exercise test: Mean endurance time for 70% of pretraining O2 consumption test improved in both groups ($p<0.01$ between groups) the difference in mean changes B-A (in sec) is 232s (41, 423)
Other, safety	ND

Ries 1986

Country:	US	$N_{centers}$:	1	RCT design:	Parallel-arm
Disease:	COPD			Setting:	Home-based

Interventions: Arm A

Exercise:	Isocapnic ventilatory muscle training with controlled flow. Intensity gradient up to as high as tolerable. (daily?) for 6 wk
Other:	All other components were common (no details)

Interventions: Arm B

Exercise:	Walking exercise training (3 times daily) intensity determined by initial testing as the one sustainable for 5 min for 6 wk
Other:	All other components were common (no details)

Patient flow

N enrolled:	?	Pre-rand exclusions (rationale)	?	N_{rand}	A= B=	10 8
Post-rand exclusions (rationale)		A= 2 dropouts due to unrelated intercurrent illnesses and 3 due to noncompliance and failure to undergo retesting; B= 1 dropout due to noncompliance		N_{analyz}	A= B=	5 7

Maximum follow-up:	At the end of the intervention (6 wk)

Population description and baseline data (SD)

Demographics:	Age (y):	A=62(6) B=67(10)	Males (%):	75%	Smokers (%):	ND
Baseline:	FEV_1 [L]:	A=1.02 (0.21) B=0.85 (0.31)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Stable COPD
Exclusion:	Significant resting or exercise hypoxemia (PaO_2 <50mmHg)

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	ND	Allocation concealment:	ND
Intention-to-treat:	no	Randomization method:	ND
Power analysis:	no	Comments:	Overall Quality C Poorly described, underpowered, focuses on physiological outcomes.

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	• differences in the change from baseline in the 12MWT were -69m (-326 to 188) • differences in the change from baseline in the endurance time at rate defined at baseline were 4.2min (-3.6, 12)
Other, safety	No

Ries 2003

Country:	USA	$N_{centers}$:	Single	RCT design:	Parallel arm
Disease:	Chronic lung disease			Setting:	Enrolled from an university rehabilitation program and maintenance at home

Interventions, experimental; Arm A

Exercise:	Experimental maintenance intervention implemented immediately after completion of the rehab program 1) weekly telephone calls and 2) monthly supervised reinforcement sessions
Other:	

Interventions: Arm B

Exercise:	"standard care" included referral back to patient's PCP with a letter recommending home based rehab and subjects invited to regular monthly alumni group meetings
Other:	

Patient flow

N enrolled:	190	Pre-rand exclusions (rationale)	N=18 (who did not agree to participate)	N_{rand}	A= B=	87 85
Post-rand exclusions (rationale)		A=12 (withdrawn=4; lung surgery=1;deceased=7) B=12(withdrawn=5; lung surgery=1;deceased=6) [comment: numbers don't add up]		N_{analyz}	A= B=	74 (12 mo) 64 (12 mo)
					A= B=	69 (24 mo) 62 (24 mo)

Maximum follow-up:	24 mo

Population description and baseline data (SD or SEM)

Demographics:	Age (y):	67.1±8.2	Males (%):	A=65% B=43%	Smokers (%):	ND
Baseline:	FEV_1 [L]:	A=1.07±0.43 B=1.14±0.42	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Other:	CRQ baseline A=103.0±18.0 B=105.9±14.3	6MWT (m) A=458.0±98.6 B=473.0±94.0	TDI A=2.9±2.4 B=2.7±2.2			
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Patients enrolled in an university pulmonary rehab program
Exclusion:	ND

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	Single	Allocation concealment:	Yes
Intention-to-treat:	No	Randomization method:	Computer program generated
Power analysis:	ND	Comments:	

Outcomes at maximum follow-up

Primary:	CRQ total at 12 mo change from baseline A= -7±3.3 p≤0.05 for time B=-10±6.7	Six min walk distance at 12 mo change from baseline (m) A=-17.9±6.3 p≤0.05 for group x time B=-42.2±36
Other, efficacy:	TDI at 12 mo change from baseline A=-2.1±0.4 p≤0.05 for time B=-1.7±0.6	
Other, safety	Mortality Total=13 died at 1 yr and 20 at 2 yr (No difference in survival between groups) Health care use: There was a significant overall reduction in maintenance group at one yr and 2 yr	
Comments:	One yr data available for 138 patients and 2 yr data for 131 patients.	

Sewell 2005

Country:	UK	$N_{centers}$:	1	RCT design:	Parallel
Disease:	COPD			Setting:	OUT

Interventions, experimental: Arm A

Exercise:	General exercise: 10 exercises for all patients (LLE + ULE, 1h aerobic and supervised circuit training), with each exercise having duration 30-120 sec; individualized intensity; 2/wk for 7wk
Other:	No

Interventions: Arm B

Exercise:	Individually targeted exercise: individualized set of personal exercises, identified through a questionnaire (COPM interview) at baseline and targeted to address activities that are discomforting to the patient (1h aerobic and supervised circuit training); individualized intensity; 2/wk for 7wk
Other:	No

Patient flow

N enrolled:	180	Pre-rand exclusions (rationale)		N_{rand}	A= B=	90 90
Post-rand exclusions (rationale)		A=exacerbation of resp illness n=23; Transport problems n=1; Deaths n=3, Decided not to continue n=4 B= exacerbation of resp illness n=15; Transport problems n=1; Deaths 3; Decided not to continue 7		N_{analyz}	A= B=	59 64

Maximum follow-up:	7 wk

Population description and baseline data (SD)

Demographics:	Age (y):	A=69.3 (8.7) B=67.3 (8.4)	Males (%):	61.1%	Smokers (%):	A=23.3 B=15.6
Baseline:	FEV_1 [L]:	A=0.93 (0.39) B=0.97 (0.45)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Other:	LTOT: A=15.6% B=8.9%					
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Stable COPD recruited from PR assessment clinic
Exclusion:	Hospital admissions or exacerbations w/in 4 wks

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	NS	Allocation concealment:	Sealed envelopes
Intention-to-treat:	NS	Randomization method:	NS
Power analysis:	>64 pts needed per/group to attain 5% signif with 80% power	Comments:	Overall Quality B Well designed, but high attrition rates.

Outcomes at maximum follow-up

Primary:	Functional activity, measured objectively with proper devices
Other, efficacy:	• CRDQ, differences in the change from baseline: Dyspnea: 0.27 (-0.22 to 0.76) Fatigue: 0.30 (-0.17 to 0.77) Emotion: -0.02 (-0.47 to 0.43) Mastery: 0.13 (-0.36 to 0.62) • Maximal exercise capacity, changes from baseline ISWT, m -3.8 (-29.1 to 21.5)
Other, safety	ND
Comments:	High attrition rates

Singh 2003

Country:	India	$N_{centers}$:	1	RCT design:	Parallel arm
Disease:	COPD, stable			Setting:	Home based

Interventions, experimental: Arm A

Exercise:	Patients instructed to walk at submaximal speed twice a day, for 30 min. Exercise performed at home or a flat track . Program lasted for 4 wks. Patients were supervised weekly to ensure compliance
Other:	Patients were taught breathing strategies (pursed lip breathing and diaphragmatic breathing); removal of secretions (postural drainage) and energy conservation and work simplification techniques

Interventions, comparator: Arm B

Exercise:	No
Other:	No

Patient flow

N enrolled:	40	Pre-rand exclusions (rationale)	0	N_{rand}	A= B=	20 20
Post-rand exclusions (rationale)		0 (None stated, assumed 0)		N_{analyz}	A= B=	20 20

Maximum follow-up:	End of intervention (at 4wk)

Population description and baseline data

Demographics:	Age (y)	59.4 (6.4)	Males (%)	32 (80)	Smokers (%)	0 (0)
Baseline:	FEV_1 [Unit]:	A=28% (7.5) B=26% (7.1)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	NS					

Patient selection criteria

Inclusion:	COPD patients with $FEV_1/FVC<0.7$ and $FEV_1<40\%$ of predicted; with dyspnea in ≥3 daily activities; who had quit smoking at least 2 months before entry; and had never participated in a PR program.
Exclusion:	Patients with right ventricular failure, unstable ischemic heart disease, O_2 saturation <20% at rest, acute exacerbation or pneumothorax

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	NS	Allocation concealed	NS
ITT	NS (but no attrition)	Randomization method	NS
Power analysis:	NS	Comments:	Overall quality C Poor overall reporting of methodological quality items

Outcomes

Primary:	Unclear
Other, efficacy:	ΔChange from baseline in: • CRDQ: CRDQ (dyspnoea)= 0.88 (0.26, 1.50) CRDQ (emotional)=0.75 (0.04, 1.46) CRDQ (fatigue)=0.84 (-0.04, 1.72) CRDQ (mastery)=0.84 (-0.11, 1.79) • 6MWT= 47.5m (-8.6, 103.6)
Other, safety	ND

Vogiatzis 2005, (RefID 28)

Country:	Greece	$N_{centers}$:	1	RCT design:	Parallel
Disease:	COPD			Setting:	Out-patient

Interventions: Arm A

Exercise:	Interval exercise Training (IE): electromagnetically braked cycle ergometers at intensity initially targeted to 100% W-peak for 30s then rest for 30 s. 45min/d, 3d/wk, 10 wks, supervised.
Other:	Breathing exercises, relaxation techniques, education, psychosocial support, nutritional intervention

Interventions: Arm B

Exercise:	Constant Load Exercise (CLE): intensity initially targeted to 60% W-peak. 45min/d, 3d/wk, 10 wks
Other:	Breathing exercises, relaxation techniques, education, psychosocial support, nutritional intervention

Patient flow

N enrolled:	19	Pre-rand exclusions (rationale)	0	N_{rand}	A= B=	10 9
Post-rand exclusions (rationale)		0		N_{analyz}	A= B=	10 9

Maximum follow-up:	End of intervention (10 wk)

Population description and baseline data (SD)

Demographics:	Age (y):	A=64 (9.5) B=67 (6)	Males (%):	84.2%	Smokers (%):	ND
Baseline:	FEV_1 [%]:	A=44% (19) B=39% (18)	PaO_2 (mmHg):	A=69 (19) B=64 (12)	$PaCO_2$ (mmHg):	A=39 (3) B=41 (3)
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Stable advanced COPD: post bronchodilator FEV_1 <50% of predicted, FEV_1/FVC <70% without significant reversibility; optimized medical treatment
Exclusion:	CVD or neuromuscular disease

Quality Assessment for RCTs:

Blinding:	Blinded assessor	Allocation concealment:	ND
Intention-to-treat:	NA (no dropouts)	Randomization method:	Stratified above and below cutoffs of 40% FEV_1 and 50W in ICET
Power analysis:	No	Comments:	Overall Quality B Small trial; Randomization unclear and the pre-randomization stratification with only 19 people is dubious; Anyway primary outcome is on biochemistry and pathology of vastus lateralis muscle.

Outcomes at maximum follow-up

Primary:	Biochemical and pathological characteristics of the vastus lateralis muscle
Other, efficacy:	Maximal exercise capacity, ICET: Significant improvement for both groups compared with baseline (p=0.04 for arm A and p=0.001 from arm B). However, A-B difference in the change from baseline was 1W (-24.7, 26.7).
Other, safety	ND

Wanke 1994 (Ref ID 1558)

Country:	Austria	$N_{centers}$:	1	RCT design:	Parallel arm
Disease:	COPD			Setting:	In-patient

Interventions: Arm A

Exercise:	Cycle ergometer training (CET) for 20 to 30 min; intensity up to 10 beats above 60% of maximal heart rate; 4d/wk for 8 wk
Other:	inspiratory muscle training (IMT) strength and endurance training with controlled flow. Daily for 8 wk.

Interventions: Arm B

Exercise:	Cycle ergometer training (CET) as in A
Other:	

Patient flow

N enrolled:	60	Pre-rand exclusions (rationale)	0	N_{rand}	A= B=	30 30
Post-rand exclusions (rationale)		18 in both arms due to exacerbation of COPD, noncompliance; "dropouts were similar in both groups", no more details.		N_{analyz}	A= B=	21 21

Maximum follow-up:	End of intervention (8 wk)

Population description and baseline data (SD)

Demographics:	Age (y):	A=55(5) B=57(6)	Males (%):	52.4 of analysed, 65% of randomised	Smokers (%):	ND
Baseline:	FEV_1 [Unit]: L	A=1.31(0.52) B=1.34(0.44)	PaO_2 mmHg	A=68.2 (9.0) B=70.5 (10.5)	$PaCO_2$ mmHg	A=42.0 (6.0) B=40.0 (6.8)
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Mild to severe COPD (FEV_1/FVC <70%; TLC >80% predicted value; change in FEV_1 after bronchodilator inhalation <15%). Ventilatory limitation of exercise: maximum HR below 2 standard deviations pf predicted max HR; exercise ventilation >80% maximum voluntary ventilation; dyspnea at maximum exercise
Exclusion:	Evidence of endocrine, orthopedic or primary cardiac disease; clinical or electrocardiograph evidence of CAD, HBT, cor pulmonale.

Quality Assessment for RCTs

Blinding:	ND	Allocation concealment:	ND
Intention-to-treat:	No	Randomization method:	ND
Power analysis:	No	Comments:	Overall Quality C 30% dropouts, no ITT analyses, they manipulate presentation of results using % improvement to claim significance.

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	Maximal exercise capacity, ICET: Both arms showed significant improvement compared to baseline in the maximum power output ($p<0.05$ in both); improvement was more impressive for arm A and was significant when the % improvements were compared between arms. However, according to the numbers in the tables, improvement in Watts was far from significant: 8.8 W (-11.1, 28.7) for the difference in the mean change from baseline.
Other, safety	ND

Wedzicha 1998 and Bestall 2003 (2 papers)

Country:	UK	$N_{centers}$:	1	RCT design:	Parallel
Disease:	COPD			Setting:	OUT for moderate/ homebased for severe

Interventions: Arm A

Exercise:	Exercise upper and lower limb training w/aerobic component (walking and cycle ergometry); intensity dyspnea limited; 2/wk for 8wk Afterwards up to 12 months: 11 exercises, 1 hr each, 1/mo, 12 mo + advice to exercise at home with frequency 3/wk to 5/wk
Other:	Education, 2/wk for 8 wk, 45 min each

Interventions: Arm B

Exercise:	No
Other:	Education, 2/wk for 8 wk, 45 min each

Patient flow

N enrolled:	138	Pre-rand exclusions (rationale)	10 declined, 66 were MRC 4 or 5, 60 were MRC 5	N_{rand}	3-4 A= B= 5 A= B=	 33 33 30 30
Post-rand exclusions (rationale)		By the end of intervention (8Wk): Grade 3-4 pts: Exercise- 3 withdrawn, 1 attended <50%. Education- 2 withdrawn, 3 attended <50%, 1 death. Grade 5 pts: Exercise- 4 withdrawn. Education - 1 withdrawn, 1 death. By the end of 1 y of follow-up in the 3-4 group: Exercise - Another 2 withdrew, another 1 death Education - Another 3 withdrew, another 3 deaths		N_{analyz}	3-4 A= B= 5 A= B=	 29 27 26 28

Maximum follow-up:	1 yr (only the moderate severity stratum)

Population description and baseline data (SD)

Demographics:		Age (y):	3-4 A=68.6 (8.9) B=68.6 (6.6) 5 A=73.9 (5.9) B=72.0 (6.1)	Males (%):	70/138 recruited (44%)	Smokers (%):	ND
Baseline:	FEV_1 [L]:	3-4 A=0.95(0.32) B=1.01(0.45) 5 A=0.87(0.41) B=0.77(0.28)		PaO_2 (mmHg):	3-4 A=67 (8) B=65 (6) 5 A=62 (12) B=64 (8)	$PaCO_2$ (mmHg):	3-4 A=44 (4) B=44 (5) 5 A=46 (7) B=45 (6)
Comorbidities:	ND						

Patient selection criteria

Inclusion:	History of COPD with FEV_1 <70% predicted, <15% reversibility due to inhaled salbutamol 400 mcg; limited level of exercise tolerance due to dyspnea; clinically stable without exacerbation within 3 wk. MRC Grades 3 and 4 were the moderate COPD stratum; MRC 5 were the severe COPD stratum. Only the moderate stratum was followed up for 1 year.
Exclusion:	Unstable angina, peripheral vascular disease, joint limiting mobility condition

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	ND	Allocation concealment:	central
Intention-to-treat:	no	Randomization method:	Computer generation, sealed envelopes
Power analysis:	no	Comments:	Overall quality B Pre-randomization stratification of patients

Outcomes at maximum follow-up

Primary:	Maximal exercise capacity (ISWT): • MRC 3-4: From baseline to end of intervention: diff in mean changes 104m (60, 148) • MRC 5: From baseline to end of intervention: diff in mean changes -4m (-31, 22) • MRC 3-4: From baseline to 1 year of follow-up there were (p=0.015 for difference between groups). The difference in the change from baseline was 68 m (11, 125) between the 2 groups

	favoring pulmonary rehabilitation with exercise training. Changes within groups NS • MRC 3-4: From end of intervention to 1y follow-up there was a steady decline in SWD during FU, F=15.97, p<0.0001 for decline in both groups For Exercise -60m (-31, -90) compared to end of intervention For Education -23m (-5, 52) compared to end of intervention.
Other, efficacy:	CRQ: • MRC 3-4: From baseline to end of intervention: diff in mean changes total of 8.9 (2.1, 15.8) • MRC 5: From baseline to end of intervention: diff in mean changes total of 0.23 (-4.9, 5.5) • MRC 3-4: Changes from baseline to 1 year in total CRQ were not statistically significant between the 2 arms p=0.112 (ANCOVA); Changes were 7 points (out of 100?) in exercise and 1 in education • MRC 3-4: Changes from end of intervention till end of follow up: Differences between gps F=6.28, p<0.016 SGRQ: • MRC 3-4: From baseline till end of intervention: diff in mean changes total of -5.5 (-10.7, 0.02) • MRC 5: From baseline till end of intervention: diff in mean changes total of 0.93 (-3.9, 5.8) • MRC 3-4: Changes from baseline till 1 year in total SGRQ were not statistically significant between the 2 arms p=0.27 (ANCOVA) • MRC 3-4: Changes from end of intervention till end of follow up: Differences between interventions p<0.05 in a repeated-measures ANCOVA, favoring by 2 units on average the exercise group.
Other, safety	Deaths: MRC 3-4: 1 died in pr, 4 in control (counting from enrollment) Deaths: MRC 5 (8ws): 0 in exercise, 1 in education

Weiner 1992 (1734)

Country:	Israel	$N_{centers}$:	1	RCT design:	3-armed parallel
Disease:	COPD			Setting:	Out-patient

Interventions: Arm A

Exercise:	Supervised general exercise reconditioning (GER). This was 20 min or cycle ergometry from low load till 50% of maximal work achieved at baseline incremental test. Then 10 min of low resistance rowing and then 15 min of strengthening exercises (LLE + ULE); 3/wk for 6mo
Other:	IMT with threshold device. From 15% up to 80% of personal maximum. + breathing exercises for 6 mo (daily?)

Interventions: Arm B

Exercise:	general exercise reconditioning (GER) alone, as in arm A for 6 mo
Other:	

Interventions: Arm C

Exercise:	control
Other:	

Patient flow

N enrolled:	36	Pre-rand exclusions (rationale)	0	N_{rand}	A= 12 B= 12 C= 12	
Post-rand exclusions (rationale)		0		N_{analyz}	A= 12 B= 12 C= 12	

Maximum follow-up:	End of intervention (6 months)

Population description and baseline data (SD)

Demographics:	Age (y):	A=67.2 (9.0) B=64.4 (10.4) C=62.3 (8.3)	Males (%):	41.7	Smokers (%):	ND
Baseline:	FEV_1 [%]:	A=33.7 (9.0) B=32.8 (3.0) C=39.5 (8.3)	PaO_2 (mmHg):	ND	$PaCO_2$ (mmHg):	ND
Comorbidities:	ND					

Patient selection criteria

Inclusion:	Spiromentric evidence of chronic airflow limitation not corrected by bronchodilator tx
Exclusion:	ND

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	ND	Allocation concealment:	ND
Intention-to-treat:	no	Randomization method:	ND
Power analysis:	no	Comments:	Overall Quality C Small trial with 3 (!) arms

Outcomes at maximum follow-up

Primary:	Unclear
Other, efficacy:	Functional exercise capacity 12MWT: arm A showed significant increase in distance walked ($p<0.0001$), B also ($p<0.001$). Control gp showed small but NS decrease. Endurance at 60% of Wmax: arm A showed stat signif improvement ($p<0.0001$) and gp B ($p<0.001$) A vs B: 12MWT 435m (173, 697) and CET, endurance at 60% of Wmax = 1.9min (-.8, 4.6) A vs C: 12MWT 550m (283, 817) Endurance at 60% of Wmax = NS
Other, safety	ND

Wijkstra 1996

Country:	Netherlands	$N_{centers}$:	Single	RCT design:	3 parallel arms
Disease:	COPD			Setting:	Home-based

Interventions: Arm A

Exercise:	Relaxation exercises; breathing training; upper limb training; target flow inspiratory muscle training; exercise trainer on a hometrainer
Other:	Twice a day for 30 min during first 3 mo; thereafter once a day for 30 min (visited by a physical therapist once a week)

Interventions: Arm B

Exercise:	Relaxation exercises; breathing training; upper limb training; target flow inspiratory muscle training; exercise trainer on a hometrainer
Other:	Twice a day for 30 min during first 3 mo; thereafter once a day for 30 min (visited by a physical therapist once a month)

Interventions: Arm C

Exercise:	None
Other:	

Patient flow

N enrolled:	45	Pre-rand exclusions (rationale)	None	N_{rand}	A= B= C=	15 15 15
Post-rand exclusions (rationale)		N=12 (9/12 dropped out during the first 12 mo. One patient in each group died, six patients dropped out because of lack of motivation or unrelated diseases - equally distributed over the 3 groups. During the last 6 mo three patients dropped out – 2 patients died in group A and B and one patient developed tumor in group C.		N_{analyz}	A= B= C=	11 12 13

Maximum follow-up:	18 mo

Population description and baseline data (SD)

Demographics:		Age (y):	A=62.3 (5.1) B=64.0 (6.2) C=61.9 (3.6)	Males (%):	A=73% B=83% C=92%	Smokers (%):	ND
Baseline:	FEV_1 [L after bronchodil]:	A=1.3 (0.4) B=1.4 (0.4) C=1.3 (0.3)		PaO_2 (mmHg):	A=8.9 (0.9) B=9.5 (1.2) C=9.6 (1.0)	$PaCO_2$ (mmHg):	A=40.0 (5.0) B=41.0 (5.0) C=41.0 (5.0)
Other:	TLC (% predicted): A=113.2 (15.9) B=119.5 (13.3) C=115.6 (18.3)						
Comorbidities:	None (pts with comorbidities excluded)						

Patient selection criteria

Inclusion:	1) postbronchodilator FEV1 <60% predicted and 2) postbronchodilator FEV1/IVC <50% (after 2 inhalations of 40 μg ipraptropium bromide)
Exclusion:	Patients with evidence of IHD, intermittent claudication, musculoskeletal disorders, or other disabling diseases were excluded.

Quality Assessment for RCTs: (yes or no) and type(s) of blinding

Blinding:	ND	Allocation concealment:	ND
Intention-to-treat:	ND	Randomization method:	Stratified randomization
Power analysis:	ND	Comments:	Stratified for FEV1 % predicted (< or ≥45% predicted), maximal work load of the bicycle ergometer test

Outcomes at maximum follow-up

<table>
<tr><td>Primary:</td><td>Bicycle Ergometer Test (results presented in figure)
Within groups analyses showed no significant changes in Wmax in groups A and B compared with their baseline value, whereas group C showed a decrease of Wmax at 12 and 18 mo. Between group analysis showed no significant differences in Wmax between the 3 groups at all time points
Within group analyses showed a significant decrease in the dyspnea score at Wmax in group A at 6, 12, and 18 mo compared with baseline. However no significant differences occurred in dyspnea score using Month 3 as a reference for the comparisons with 6, 12, and 18 mo in all groups. At all time points the differences in dyspnea score between the three groups were not statistically significant.
<table>
<tr><td>T1</td><td>Borg score of dyspnea</td><td>7.1±2.7</td><td>-2.6±0.2
Sig with baseline</td></tr>
<tr><td>T2</td><td></td><td>5.1±2.2</td><td>-0.5±0.1</td></tr>
<tr><td>C</td><td></td><td>6.4±2.0</td><td>-0.8±0.4</td></tr>
</table>
Six minute walking distance
Within group analyses showed that 6 min walking distance in group C decreased after 12 mo ($p<0.05$) and after 18 mo ($p<0.01$) compared with baseline, whereas no significant changes occurred in groups A and B. At none of the time points did significant differences between the 3 groups occurred
<table>
<tr><td>T1</td><td>6 min walk</td><td>438±9.0</td><td>+12.0±20</td></tr>
<tr><td>T2</td><td></td><td>466±26.0</td><td>-16.0±3.0</td></tr>
<tr><td>C</td><td></td><td>462±8.0</td><td>-12±21</td></tr>
</table>
Strength and Endurance capacity of the Inspiratory muscles
Group A showed a significant increase of both PIP and endurance capacity at 3 and 12 mo compared to baseline; Group B and C showed no significant changes. Between group analysis showed no significant differences in PIP and endurance capacity between the 3 groups at any time point</td></tr>
<tr><td>Other, efficacy:</td><td></td></tr>
<tr><td>Other, safety</td><td></td></tr>
<tr><td>Comments:</td><td></td></tr>
</table>