Does pulmonary rehabilitation work in clinical practice? A review on selection and dropout in randomized controlled trials on pulmonary rehabilitation

Bodil Bjoernshave1
Jens Korsgaard2
Claus Vinther Nielsen3
1 Medical Department, Horsens Regional Hospital, Denmark;
2 Aalborg Hospital Science and Innovation Centre, 3 Department of Clinical Social Medicine and Rehabilitation, Institute of Public Health, Aarhus University, Denmark

Aim: To analyze randomized controlled trials (RCTs) on pulmonary rehabilitation (PR) to determine whether the patients who complete PR form a representative subset of the chronic obstructive pulmonary disease (COPD) target population and to discuss what impact this may have for the generalizability and implementation of PR in practice.

Material and methods: A review of 26 RCTs included in a Cochrane Review 2007. We analyzed the selection at three different levels: 1) sampling; 2) inclusion and exclusion; 3) and dropout.

Results: Of 26 studies only 3 (12%) described the sampling as the number of patients contacted. In these studies 28% completed PR. In all we found, that 75% of the patients suitable for PR programs were omitted due to sampling exclusion and dropout. Most of the study populations are not representative of the target population.

Conclusion: The RCTs selected for the Cochrane review gave sparse information about the sampling procedure. The demand for high internal validity in studies on PR reduced their external validity. The patients completing PR programs in RCTs were not drawn from a representative subset of the target population. The ability to draw conclusions relevant to clinical practice from the results of the RCTs on PR is impaired.

Keywords: COPD, rehabilitation, selection, dropout, external validity

Introduction

The primary goal of pulmonary rehabilitation (PR) is to restore the patients to the highest possible level of independent function, and the target population are patients with stable chronic obstructive pulmonary disease (COPD).1–6 PR evidently seems to benefit the patients in terms of quality of life, functional capacity, symptom relief, and reductions in exacerbation of the condition and in the number of days in hospital. PR is therefore recommended in all influential guidelines based on grade A.2,3,6,7

The concept of PR rests on a large number of randomized controlled trials (RCTs) and is defined 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”.4 Patients must accordingly exercise with a certain intensity, duration and frequency, and they must be well-informed about their disease through interventions such as patient education, together with being taught...
self-help and coping strategies. In addition, the patients must perform medication and breathing techniques correctly, and they must stop smoking.

Since the 1980s, at least 600 controlled trials on COPD and rehabilitation have been published. During this period, the PR concept has expanded, specific components have been developed, and techniques have steadily become more comprehensive and sophisticated.8-10

PR is recommended in Denmark and it has therefore been suggested that any COPD patient who is registered with severe dyspnea measured with Medical Research Council Dyspnea Scale MRC grade ≥3 should be offered PR.12 However, a report from the Danish National Indicator Project13 documents that approximately only 60% (CI: 59–62) of patients registered with MRC grade ≥3 are offered rehabilitation. Also there is no documentation of number of patients that complete a program or the effects of the rehabilitation. This shows that we do not know who actually completes PR nor who will benefit from it, when it is to be implemented in clinical practice.12

In general, COPD patients have extensive rehabilitation needs because their disease imposes major restrictions on their everyday life;2,3,4,6 however, poor adherence is common in daily practice as some patients fail to attend programs and others drop out.3,4,6,7,14-18 Our experience is that in order to optimize the resources used on PR, patients are selected so that those who are deemed to have the ability and motivation to complete a PR program are more likely to be chosen for participation than patients with poor motivation. This may entail understandable, but ethically inappropriate inequality in access to health care.

The effect of PR is well documented in RCTs and its rationale can therefore hardly be questioned as RCTs are considered to represent the most scientifically rigorous method of hypothesis testing in order to provide the best evidence.7,19 RCTs must satisfy strict quality criteria and explicit standards regarding patient selection. However, RCTs on PR often fail to adequately discuss their external validity, ie, the ability to “produce unbiased inferences regarding the target population”.19 We may therefore justifiably question whether the reported effects can indeed be generalized to the target population.

Selection in RCTs may take place at three different levels: 1) sampling; 2) inclusion and exclusion; and 3) dropout. At the sampling level, a number of COPD patients are selected among all subjects within a particular population.19 The selection criteria used prior to randomization, ie, when some patients are contacted for screening, and others are not, need to be explicit. At the inclusion/exclusion level, criteria are defined to establish the study population and to homogenize the intervention group and the control group. This level faces the risk that the patients included differ in certain aspects from those who are not included.

At these first two levels, selection is a matter of the investigator’s choice and ideally, information about all patients who are not included must be registered in order to optimize the external validity.

At the third and final level, dropout may contribute to a weakening of the internal validity; ie, dropout may cause that the observed differences between the compared groups, may not be attributed only to the hypothesized effect under investigation.19

The challenges involved in achieving high internal and external validity makes selection of patients for PR a pertinent issue. Patients, clinicians and decision-makers need clear messages about the evidence of PR to accept its widespread application and to ensure that scarce resources can be used to good effect. Dropout is usually well-described in RCTs on PR, but information on selection performed during sampling, ie, before randomization, remains sparse. There would therefore seem to be some room for strengthening the discussion of the validity of RCTs in general, and of their external validity in particular by examining pre-randomization selection issues at greater depth.

The aim of the present study is to analyze RCTs on PR to determine whether the patients who complete PR form a representative subset of the COPD target population and to discuss what impact this may have for the generalizability and implementation of PR in practice.

Material and methods
A literature review of the RCTs originally included in the Cochrane review10 published in 2007 endeavoured “to establish the influence and the effect size of pulmonary rehabilitation on health related quality of life, functional capacity in patients with COPD”. The review stated that “Rehabilitation relieves dyspnea and fatigue, improves emotional function and enhances patients’ sense of control over their condition”. This review was chosen for this present analysis because it strived to comprehensively identify and synthesize all the literature on PR, and it is in general, well-reputed and often cited.

The Cochrane review included a total of 31 RCTs of which 26 full-text English language versions were examined. Five studies were not examined; three studies were only available as English abstracts as the articles were published in Spanish,20 French,21 and Chinese.22 Boxall23 was not used in the form, ie, congress abstract, in which it was used in the
Cochrane review, it was instead published as an article in 2006. Chlumsky 2001 was not found. Casaburi26 was not included in this analysis as its focus on testosterone supplementation was deemed irrelevant to the present purpose.

The 26 RCTs were analyzed with regard to their description of sampling, inclusion and exclusion criteria, and dropout illustrated in Figure 1.

Correlation analyses were performed to examine possible associations between selection criteria, disease events, the rehabilitation program, and the number of patients left out.

Results
The results of the analysis are described in three sections corresponding to the three levels: 1. sampling, 2. inclusion and exclusion, and 3. dropout.

1. Sampling
Only three (12%) out of 26 studies described the number of patients contacted (Table 1).27–29 In these studies, a total of 322 patients were contacted and out of those 151 (47%) were left out without being screened.

These three studies did not differ from the other 23 studies18,24,30–50 in relation to the number of patients randomized, inclusion or exclusion criteria, dropout, or lung function.

Details regarding the circumstances under which the studies were carried out revealed that Jones and colleagues28 searched computerized records to identify regular attenders at their clinic. Bendstrup and colleagues27 invited patients who were chosen from hospital records, however, the total number of relevant records was not described. Ringbaek and colleagues29 contacted the patients with moderate COPD, who were recruited from an outpatient clinic during a six-month period, however, the total number of relevant records was not described.

In brief, the size of the total COPD population, which is relevant for the external validity of the studies reviewed, is largely unknown and only a few studies were explicit about the characteristics of the populations from which the study populations were drawn. Nearly half of the patients contacted were not offered screening, and only one third of the patients contacted actually completed the PR program.

Figure 1 The different levels of selection.
Table 1 Studies with a description of the selection procedure from the number of patients contacted to the number of patients completed (3/26 articles)

<table>
<thead>
<tr>
<th>Ref number</th>
<th>Study/aim</th>
<th>Number of patients contacted</th>
<th>Number of patients screened</th>
<th>Left out from number of patients contacted to number of patients screened (%)</th>
<th>Number of patients randomized</th>
<th>Number of patients left out from number of patients screened to number of patients randomized (%)</th>
<th>Number of dropouts after randomization (%)</th>
<th>Patients completed (%) out of number of patients a) contacted b) screened c) randomized</th>
<th>Number of exclusion criteria a) disease-specific b) others</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>Jones 1985 Examine the effectiveness of inspiratory muscle training</td>
<td>52</td>
<td>38</td>
<td>14/52 (27)</td>
<td>30</td>
<td>8/38 (21)</td>
<td>9/30 (30) (5: illness, 1: lack of time, 1: social breakdown, 1: stroke, 1: lost interest)</td>
<td>a) 21/52 (40) b) 21/38 (55) c) 21/30 (70)</td>
<td>2/4</td>
</tr>
<tr>
<td>27</td>
<td>Bendstrup 1997 Investigate the effects on activities of daily living QoL and exercise tolerance of an out-patient rehabilitation</td>
<td>140</td>
<td>85</td>
<td>55/140 (39)</td>
<td>42</td>
<td>43/85 (51)</td>
<td>10/42 (24) (1: died, 1: acute abdominal surgery, 1: AMI, 7: did not want to participate)</td>
<td>a) 32/140 (23) b) 32/85 (38) c) 32/42 (76)</td>
<td>3/3</td>
</tr>
<tr>
<td>29</td>
<td>Ringbaek 2000 Investigate feasibility, effect and economic aspects of a rehabilitation programme consisting of two sessions a week for 8 weeks</td>
<td>130</td>
<td>48</td>
<td>82/130 (63)</td>
<td>45</td>
<td>3/48 (6)</td>
<td>7/45 (16) (3: exacerbation, 2: joint or back pain, 1: lack of time, 1: failed to appear)</td>
<td>a) 38/130 (29) b) 38/48 (79) c) 38/45 (84)</td>
<td>4/4</td>
</tr>
</tbody>
</table>

Total N/ (%) 322 171 (53%) 151 (47%) 117 54 (44%) 26/117 (22%) a) 91/322 (28%) b) 91/171 (53%) c) 91/117 (78%)
2. Inclusion and exclusion

Information regarding the number of patients screened and the number of patients de-selected is mainly due to inclusion and exclusion criteria, which was stated in the three above-mentioned studies (Table 1) and in five other studies (Table 2).\textsuperscript{18,33,36,39,40}

In all, eight studies screened a total of 1,040 patients of whom 406 (39%) were de-selected before randomization, and the de-selection percentage ranged from 8%–64%.

Reasons for leaving out some patients were, eg, those who lived too far away or whose social circumstances affected their ability to complete the program.\textsuperscript{18} Only patients who had the ability to travel independently to a physiotherapy practice and the motivation to improve self-care were chosen for participation.\textsuperscript{33}

A study recruited patients by means of announcements and some were screened over the telephone, however, neither the exact number of responders nor the number of telephone screenings were reported.\textsuperscript{36}

In general only stable patients were randomized. The majority of the studies (18 out of 26) were not explicit regarding the number of patients de-selected due to exclusion criteria.

The correlation between the proportion of patients left out from screening to randomization and the total number of patients screened showed a correlation, the Spearman’s correlation coefficient (0.643) tended to be significant ($P = 0.086$).

This indicates that the higher the number of patients screened the stricter the screening procedures were. There was no statistically significant correlation between the number of patients left out from screening to randomization and the number of diagnosis-specific inclusion and exclusion criteria ($P = 0.9$) nor the number of non-disease-specific inclusion and exclusion criteria ($P = 0.3$).

The exclusion criteria varied. Up to 6 pulmonary disease-specific exclusion criteria were used in the studies, eg, lung function, dyspnea, arterial blood gases, oxygen saturation, and smoking. One to seven non-pulmonary disease-specific criteria were used in the studies, eg, ischemic heart disease, cognitive impairment, musculoskeletal disorders, social circumstances, transport difficulties, or language barriers.\textsuperscript{18,33}

To summarize, the wide range (6%–64%) of patients left out from screening to randomization could not be explained by the number of explicit exclusion criteria. The analysis showed that a total of 8 (31%)\textsuperscript{18,27–29,33,36,39,40} of the 26 articles documented the sampling procedure from screening to randomization. The rate of patients completing a PR program out of the total number of patients screened ranged from 32% to 100%.

3. Dropout

Dropout was described in all 26 studies (Table 3).\textsuperscript{18,24,27–34,36–46,48–52} Dropout ranged from 0% to 36% in the intervention groups (mean = 17[CI: 12–22]) and from 0% to 54% in the control groups (mean = 15 [CI:9–22]). We found no differences in dropout between the intervention groups and the control groups ($P = 0.4$), and the correlation analysis showed a statistically significant, positive correlation (Spearman 0.72, $P = 0.00$).

The reasons for dropout can roughly be divided into two categories: “disease-related reasons,” eg, exacerbations, illness and death, and “other reasons”, eg, lack of time, motivation or cooperation (ie, patients did not want to participate, travelling difficulties, the PR-programme was too hard). In 17 out of 26 studies, “other reasons” for dropout were described.

None of the studies discussed the possibility that dropout may cause misclassification,\textsuperscript{33} or the direction of possibly biased estimates.

Altogether we found, when the size of the unknown total population was ignored, that on average, three quarters of the patients most likely suitable for PR seemed to have been de-selected, probably in a biased way, due to sampling, exclusion criteria, and dropout. None of the studies discussed generalizability and applicability.

Discussion

The present study aimed to determine whether the patients who complete PR form a representative subset of the target population. This study details aspects of patient selection for RCTs based on the sampling procedures described in the RCTs on PR included in a Cochrane review. The main result of the study is that most RCT study populations are not sufficiently representative of the COPD target population. This seriously affects the external validity of these studies and may inhibit the implementation and effects of PR in clinical practice.

The discussion is divided into four parts. The first part concerns the target population and the following three parts discuss the different levels of selection as illustrated in Figure 1.

The target population

For the sake of generalizability the study population must be drawn from a representative subset of the target population,\textsuperscript{53} which, for PR, would comprise patients diagnosed with stable COPD. However, the target population is not easily determined as the COPD prevalence is generally difficult to estimate.\textsuperscript{54–56} Firstly, population-based estimates
**Table 2** Studies with description of selection procedure from number of patients screened too number of patients completed (5/26 studies)

<table>
<thead>
<tr>
<th>Ref number</th>
<th>Study/aim</th>
<th>Number of patients screened</th>
<th>Number of patients randomized</th>
<th>Number (%) left out from screening to randomization</th>
<th>Dropout/ randomized (%)</th>
<th>Dropout reasons</th>
<th>Dropout % intervention/ control</th>
<th>Patients completed/ a) screened (%)</th>
<th>Patients completed/ b) randomized (%)</th>
<th>Number exclusion criteria disease-specific/others</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Goldstein 1994</td>
<td>244</td>
<td>89</td>
<td>155/244 (64)</td>
<td>11/89 (12)</td>
<td>Inconvenience of travelling (1), anger (1), anxiety (2), exacerbation (3), heart disease (1), arthritis (1), smoking (2)</td>
<td>16/9</td>
<td>a)78/244 (32)</td>
<td>b) 78/89 (88)</td>
<td>4/7</td>
</tr>
<tr>
<td>33</td>
<td>Cambach 1997</td>
<td>130</td>
<td>89</td>
<td>41/130 (32)</td>
<td>23/89 (26)</td>
<td>Intensity of programme too high (9), illness (5), moved away (1), refused to take part (1), full-time job (3), could not be tested within three weeks (2), did not complete final test (1), broke arm (1)</td>
<td>20/33</td>
<td>a) 66/130 (51)</td>
<td>b) 66/89 (74)</td>
<td>6/5</td>
</tr>
<tr>
<td>36</td>
<td>Emery 1998</td>
<td>92</td>
<td>79</td>
<td>13/92 (14)</td>
<td>6/79 (8)</td>
<td>Illness (4), transport difficulties (2)</td>
<td>13/4</td>
<td>a) 73/92 (79)</td>
<td>b) 73/79 (92)</td>
<td>3/5</td>
</tr>
<tr>
<td>40</td>
<td>Güell 1998</td>
<td>65</td>
<td>60</td>
<td>5/60(8)</td>
<td>none</td>
<td>None</td>
<td>0/0</td>
<td>a) 60/65 (92)</td>
<td>b) 60/60 (100)</td>
<td>4/2</td>
</tr>
<tr>
<td>39</td>
<td>Griffiths 2000</td>
<td>338</td>
<td>200</td>
<td>138/338 (41)</td>
<td>16/200(8)</td>
<td>Death (4), missing data (12)</td>
<td>10/6</td>
<td>a) 184/338 (54)</td>
<td>b) 184/200 (92)</td>
<td>3/4</td>
</tr>
</tbody>
</table>

**Total table 2** 869 517 352/869 (41) 56/517 (11) 1 a) 461/869 (53) b) 461/517 (89) 3/4

**Total table 1 + 2** 1040 634 406/1040 (39) 82/634 (13) 1 a) 552/1040 (53) b) 552/634 (87)
Table 3  Studies were only the number of patients randomized were described, number of exclusions criteria and dropout (17/26 studies)

<table>
<thead>
<tr>
<th>Ref number</th>
<th>Study/aim</th>
<th>Randomized/completed (%)</th>
<th>Exclusion disease-specific/others</th>
<th>Dropout (%) (Intervention/control %)</th>
<th>Reasons dropout</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>McGavin 1977 Evaluate a training scheme carried out by the patients unsupervised at home</td>
<td>28/24 (86)</td>
<td>3/4</td>
<td>4/28 (14) (29/0)</td>
<td>lack of enthusiasm (2), depressive (1), death (1)</td>
</tr>
<tr>
<td>35</td>
<td>Cockcroft 1981 Evaluate the effects of exercise training in men with chronic respiratory disability</td>
<td>39/34 (87)</td>
<td>1/2</td>
<td>5/39 (13) (5/20)</td>
<td>deterioration in condition (2) stroke (1) abroad (1) domestic problems (1)</td>
</tr>
<tr>
<td>31</td>
<td>Booker 1984 Longterm RCT to investigate the subjective and objective effects of progressive exercise training in patients with chronic airflow limitation</td>
<td>128/102 (80)</td>
<td>1/1</td>
<td>26/128 (20) (23/15)</td>
<td>no reasons described</td>
</tr>
<tr>
<td>32</td>
<td>Busch 1988 Effects of a 18 weeks home exercise program on physical work capacity and dyspnea</td>
<td>20/14 (70)</td>
<td>1/3</td>
<td>6/20 (30) (30/30)</td>
<td>death (1), exercising of own volition (2), did not perform exercise regularly (3)</td>
</tr>
<tr>
<td>42</td>
<td>Lake 1990 Evaluate the benefit of upper-limb exercise training alone and in combination with walking training</td>
<td>28/26 (93)</td>
<td>4/7</td>
<td>2/28 (7) (7/7)</td>
<td>infection (1) cerebral attack (1)</td>
</tr>
<tr>
<td>45</td>
<td>Simpson 1992 Determine whether specific muscle training techniques are helpful</td>
<td>34/28 (82)</td>
<td>2/3</td>
<td>6/34 (18) (18/17)</td>
<td>infection (1), change in treatment (2) unknown reasons (3)</td>
</tr>
<tr>
<td>49</td>
<td>Weiner 1992 Effect of specific inspiratory muscle training combined with exercise reconditioning for six months</td>
<td>36/36 (100)</td>
<td>1/3</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Reardon 1994 Evaluate the effect of outpatient pulmonary rehabilitation on dyspnea</td>
<td>20/20 (100)</td>
<td>3/1</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Wijkstra 1994 Investigate the effect of home rehabilitation programme on QoL and exercise tolerance</td>
<td>45/43 (96)</td>
<td>4/4</td>
<td>2/45 (4) (7/0)</td>
<td>cerebral tumor (1), arthrosis (1)</td>
</tr>
<tr>
<td>34</td>
<td>Clark 1996 Investigate physiological effects of a 12 week programme of conditioning of peripheral muscle</td>
<td>48/48 (100)</td>
<td>1/0</td>
<td>none</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Ref number</th>
<th>Study/aim</th>
<th>Randomized/completed (%)</th>
<th>Exclusion disease-specific/others</th>
<th>Dropout (%) (Intervention/Control %)</th>
<th>Reasons dropout</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>Strijbos 1996 Effects of a 12 weeks outpatient pulmonary rehabilitation compared with home-based program</td>
<td>50/45 (90)</td>
<td>6/3</td>
<td>5/50 (10) (20/7)</td>
<td>lack of motivation (2), death (2), cancer (1)</td>
</tr>
<tr>
<td>37</td>
<td>Engstrøm 1999 To examine long-term effects of outpatients rehabilitation</td>
<td>55/50 (91)</td>
<td>5/3</td>
<td>5/55 (9) (7/11)</td>
<td>death (3), heart disease (1), did not complete (1)</td>
</tr>
<tr>
<td>30</td>
<td>Behnke 2000 Examine home-based walking training</td>
<td>46/30 (65)</td>
<td>2/3</td>
<td>16/46 (35) (34/34)</td>
<td>death (2), exacerbation (4), lack of motivation (6), unrelated diseases (4)</td>
</tr>
<tr>
<td>48</td>
<td>Troosters 2000 Investigate short- and long-term effects of 6 months programme</td>
<td>62/100 (62)</td>
<td>2/5</td>
<td>38/100 (38) (32/44)</td>
<td>refused follow-up (33), death (5)</td>
</tr>
<tr>
<td>41</td>
<td>Hernandez 2000 Investigate the effectiveness of a home-based program of exercise training</td>
<td>60/37 (62)</td>
<td>4/6</td>
<td>23/6 (38) (33/43)</td>
<td>(lack of cooperation (13), exacerbation (8), cerebral accident (1), cancer(1)</td>
</tr>
<tr>
<td>38</td>
<td>Finnity 2001 Assess the effectiveness of outpatient-based PR</td>
<td>100/55 (65)</td>
<td>3/7</td>
<td>45/100 (45) (36/54)</td>
<td>failed to attend (27), failed to continue (18), reasons not described</td>
</tr>
<tr>
<td>46</td>
<td>Singh 2003 Evaluate the effect of domiciliary PR</td>
<td>40/40 (100)</td>
<td>6/4</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Boxall 2003/2005 Evaluate a 12 week home-based PR</td>
<td>60/46 (77)</td>
<td>2/3</td>
<td>14/60 (23) (23/23)</td>
<td>hip fracture (1), exacerbation (1), exercises to hard (3), cancer (1), death (1), reasons not described (7)</td>
</tr>
</tbody>
</table>

of COPD prevalence are complicated by the variety of tools and definitions used to describe COPD. Secondly, COPD terminology is inconsistent and widely accepted diagnostic standards are lacking, therefore COPD coding is insufficient and COPD data often inaccurate. Thirdly, the method by which prevalence is estimated (expert opinion, patient reporting, symptom reporting or measurement by spirometry) influences the reported prevalence estimates. Furthermore patients who participate in RCTs on PR are selected among patients already in clinical practice, eg, clinic, hospital or by GP, and not amongst patients identified from prevalence studies. For these reasons the exact target population in RCTs remains more or less unknown. None of the studies in the Cochrane review discussed whether the study population was a representative subset of the target population.

Pre-randomization selection
Most studies failed to describe which patient were selected and those who were not; so information is lacking on the number of patients left out and their characteristics. Only three (12%) of the included studies were explicit about who were contacted. We have estimated that almost half (47%) of the patients were left out before screening.

Of particular concern is that none of the studies were explicit about the selection of cases and the bias this may cause. We therefore cannot be sure that the study populations were representative subsets of the target populations studied in relation to the effect of PR. This obviously leaves some severe difficulties in generalizing the findings, both in terms of capability to complete, and on the effects of participating in PR.
Representativeness should be ensured by randomly selecting the study population so that, ideally, all patients relevant for a study would have the same chance of enrolment. This requires that the investigator controls the target population and pays attention to any difference between the patients who were selected and those who were not. During this selection process, patients who, eg, were deemed not to have the ability to complete the programs, who lived too far away, or who had difficult social circumstances would be at risk of being left out. Such left out patients would likely differ from those who were included in relation to important variables, eg, incriminating psychosocial situation.

Inclusion and exclusion criteria

Only in 8 of the 26 studies was number of patients deselected during screening explicit (Tables 1 and 2). Approximately one third of the patients were left out. No correlation was found between the number of patients lost from screening to randomization and the number of inclusion and exclusion criteria; therefore the number of patients left out cannot be directly explained with reference to these criteria. The study populations must be pathophysiologically uniform to optimize study power and therefore strict inclusion and exclusion criteria must be used. This premise is obvious but there might still be a risk that nonexplicit criteria are responsible for the number of patients left out. This complicates the implementation of PR, as it is not known whether the findings can be applied to the population which is to benefit from PR.

We assume that the reviewed RCTs a priori recruited the most motivated patients to maximize completion and compliance with the intervention. Our analysis revealed that those who declined the intervention stated a variety of reasons, eg, skepticism as to their ability to attend and to adhere to study requirements, problems of transportation. Including smokers in PR programs is often conditional on their participation in a smoking cessation program, and smokers are therefore less likely to participate.

In summary there is a risk that patients selected for participation differed from the deselected patients in relation to social factors, co-morbidity and their general health status. The selection may not have been truly random, and therefore skewed by explicit as well as non-explicit choices. Implicit criteria hinder the possibility that effects, as documented in RCTs, are to be reached, when PR is implemented in the target population.

Dropout

In accordance with the Helsinki declaration, dropout was described in all 26 studies. Some dropout should be expected in PR due to the natural history of COPD and therefore selection is usually carried out to reduce dropout.

Our analyses showed huge variation in dropout rates (ranging from 0% to approximately 50%). Moreover, a strong positive correlation was observed between dropout in the intervention groups and in the control groups, which means that only a small part of the dropout can be explained by the intervention. Contrary to expectation, the number of explicit exclusion criteria and the duration of the intervention did not explain the wide variation in dropout.

The patients’ own views and experiences may influence dropout, (eg, if the patients think that the programs are either, too hard, too demanding, too difficult or too easy, unlikely to be helpful, and a waste of time), in which case they will be more likely to drop out. If the patients do not feel comfortable or unsafe and not at ease with the health professionals, or with the other patients, they may also tend to drop out. Whereas the support of family and peers may enhance patient adherence. Besides, the patients’ perception of their illness and its management may have an influence. If the rehabilitation program makes common sense, in relation to the patients’ personal beliefs about their illness, completion may be more likely to be achieved. Empathic understanding and the practitioner relationship may also have some impact.

In general, the differences between completers and dropouts are unknown and it is difficult to determine, how dropout has affected the results.

In summary

The question raised in this study is whether the findings of the RCTs and the recommendations that emanate from the Cochrane review apply to the entire PR target population. Our analyses show uncertainty regarding the representativeness of the COPD patients who completed the PR programs.

If PR, based on RCTs, is only applicable to a limited subset of the target population, it may lead to social inequality. Patients with severe disease may be too physically impaired to actively participate in and benefit from a program. Patients with a mild disease and minimal limitations may not benefit from the program because of a lack in perceived need and motivation. Patients at different stages of COPD have different needs, as the management of mild and moderate COPD involves the avoidance of risk factors in order to prevent the progression of the disease and pharmacotherapy
to control symptoms. Severe and very severe COPD require integration of several different disciplines, a variety of treatment approaches, and continued patient support.1

PR is a recommended standard of care that encourages patients to undertake their own health care and to become less dependent on health professionals and expensive medical resources. PR focuses on reducing disability from the disease despite the severity. For these reasons, it is problematic that the evidence for PR rests on studies where the study population is not representative, eg, that COPD patients with comorbidity and different social circumstances are excluded.

**Conclusion**

In conclusion, the RCTs selected for the Cochrane review comprised information about the included patients and drop-outs, though not about the sampling procedure. The internal validity was assessed in order to examine the relationship between the intervention and the treatment effect. The demand for high internal validity in studies of the effect of PR reduced their external validity, which was not assessed.

The patients completing PR programs in RCTs were not drawn from a representative subset of the target population. A number of criteria which were not explicit were used during the sampling. The studies did not meet the ideal demands for representativeness, which should be obtained by random sampling, and for equality of opportunity to participate for all patients. This impairs our ability to draw general conclusions relevant to clinical practice from the results of the RCTs on PR.

To strengthen the external validity from studies on PR, there is an extensive need for explicitness in all levels of patient selection. Studies on PR in nonselected target populations drawn from prevalence studies are needed in order to determine the number of completers and the effects gained. In addition, studies focusing on potential differences between COPD patients who complete, drop out or de-select rehabilitation are relevant when discussing how PR can be implemented to a larger proportion of the target population.

**Disclosures**

The authors have no conflicts of interest relevant to the outcome of this work.

**References**


