Effectiveness of incentive spirometry in patients following thoracotomy and lung resection including those at high risk for developing pulmonary complications

Paula Agostini, Babu Naidu, Hayley Cieslik, Richard Steyn, Pala Babu Rajesh, Ehab Bishay, Maninder Singh Kalkat, Sally Singh

ABSTRACT

Background Following thoracotomy, patients frequently receive routine respiratory physiotherapy which may include incentive spirometry, a breathing technique characterised by deep breathing performed through a device offering visual feedback. This type of physiotherapy is recommended and considered important in the care of thoracic surgery patients, but high quality evidence for specific interventions such as incentive spirometry remains lacking.

Methods 180 patients undergoing thoracotomy and lung resection participated in a prospective single-blind randomised controlled trial. All patients received postoperative breathing exercises, airway clearance and early mobilisation; the control group performed thoracic expansion exercises and the intervention group performed incentive spirometry.

Results No difference was observed between the intervention and control groups in the mean drop in forced expiratory volume in 1 s on postoperative day 4 (40% vs 41%, 95% CI −5.3% to 4.2%, p=0.817), the frequency of postoperative pulmonary complications (PPC) (12.5% vs 15%, 95% CI 7.9% to 12.9%, p=0.803) or in any other secondary outcome measure. A high-risk subgroup (defined by ≥2 independent risk factors; age ≥75 years, American Society of Anaesthesiologists score ≥3, chronic obstructive pulmonary disease (COPD), smoking status, body mass index ≥30) also demonstrated no difference in outcomes, although a larger difference in the frequency of PPC was observed (14% vs 23%) with 95% CIs indicating possible benefit of intervention (−7.4% to 2.6%).

Conclusions Incentive spirometry did not improve overall recovery of lung function, frequency of PPC or length of stay. For patients at higher risk for the development of PPC, in particular those with COPD or current/recent ex-smokers, there were larger observed actual differences in the frequency of PPC in favour of the intervention, indicating that investigations regarding the physiotherapy management of these patients need to be developed further.

BACKGROUND

Following thoracotomy, patients frequently receive routine respiratory physiotherapy treatments in order to prevent the type of postoperative pulmonary complications (PPC) associated with significant clinical and economic impact. Incentive spirometry is sometimes included by physiotherapists in the treatment of these patients, and is a breathing technique where deep breathing exercises...
are performed through a device offering visual feedback, both in terms of inspired flow and/or volume. Deep breathing is thought to re-expand areas of collapsed lung postoperatively by stretching the tissue, and to mobilise secretions. The addition of visual feedback is thought to improve breathing technique and improve patient motivation.

Routine physiotherapy provision following thoracic surgery has long been advocated, with the general belief by thoracic surgeons that it is beneficial. However, to date, no randomised studies support the provision of targeted respiratory physiotherapy over standardised care (including early mobilisation) following major thoracic surgery and there is little evidence specifically concerning the use of incentive spirometry with only two small comparative studies demonstrating no benefit in either recovery of lung function or frequency of PPC.

Reduction in pulmonary morbidity and costs for thoracic surgery patients receiving respiratory physiotherapy and relatively vigorous exercise have been observed, but these studies were non-randomised and the need remains for further higher quality research to confirm such benefit and to identify exactly which physiotherapy interventions are most useful and in which patients (based on risk). This is of increasing importance as new European and British guidelines permit surgery to be considered in increasingly higher risk patients.

The primary aim of this study was to examine the effectiveness of incentive spirometry compared with deep breathing exercises in patients undergoing thoracotomy and lung resection in a randomised controlled trial (RCT). The secondary aim was to observe the outcomes of similar patients in subgroups defined by risk factors for the development of PPC.

**METHODS**

**Design**

A prospective open parallel design was employed in this pragmatic single-blind RCT conducted between October 2008 and October 2010. Eligible patients were randomly allocated to either control or intervention groups using block randomisation after giving written consent.

**Participants**

Participants were identified from the accessible population of patients undergoing planned thoracotomy and lung resection at a tertiary regional thoracic centre. Patients were considered to be eligible for study after preoperative screening if they were scheduled for a planned thoracotomy and lung resection and aged 18 or over. Patients were excluded if undergoing emergency thoracotomy, procedures involving the mediastinum and chest wall or planned lung resection via video-assisted thoroscopic surgery, or if they were immobile preoperatively or unable to perform preoperative spirometry or the allocated breathing exercise. Decisions regarding patient operability and resectability were informed by national guidelines. Postoperative pain control was achieved by continuous thoracic epidural intrathecal morphine and/or intercostal blocks or systemic opioids. On postoperative day (POD) 1, patients were reviewed during initial physiotherapy assessment and those considered suitable were randomised using block randomisation in blocks of four. This prevented randomisation of patients whose ability to participate might have changed following anaesthesia or who had undergone different surgical procedures to those planned.

**Intervention**

Thoracic expansion exercises are deep breathing exercises emphasising active inspiration with a breath hold before a passive expiration. Incentive spirometry is deep breathing performed through a visual feedback device encouraging maximal inspiration and including a breath hold; the Coach 2 incentive spirometry device (Medimark Europe, Grenoble, France) was used in this study. On the morning of POD 1, all subjects were assessed by a physiotherapist working exclusively in the area of thoracic surgery; control group subjects then commenced thoracic expansion exercises and intervention group subjects started incentive spirometry with the Coach 2 device. Both groups received a supervised session (including 10 repetitions) of their allocated breathing exercise in the morning and then again in the afternoon of POD 1. From POD 2, subjects received one supervised session of their allocated breathing exercise unless a higher frequency was deemed necessary by the attending physiotherapist based on clinical need. From the first supervised session, all subjects were encouraged to repeat their allocated breathing exercise independently, 10 times hourly (during waking hours). Treatment also included supported coughing, early mobilisation and active shoulder exercises in accordance with current UK practice. Subjects continued to receive their allocated breathing exercise until hospital discharge. If any patient’s condition required escalation of physiotherapy treatments in terms of frequency or intensity, this was not withheld.

**Outcome measures**

The primary outcome measure observed was percentage drop in forced expiratory volume in 1 s (FEV₁) on POD 4 (POD 4 FEV₁) expressed as a percentage of the preoperative FEV₁ value. This time point was chosen as it is during the initial period that substantial loss in lung volume occurs and recovery might be best affected by therapies such as incentive spirometry. FEV₁ was measured with Vitalograph spirometry equipment (Vitalograph, Buckingham, UK) by a blinded assessor; physiotherapy interventions were not provided by this assessor nor was the assessor aware of group allocation.

Secondary outcomes included the mean percentage predicted postoperative (ppo) FEV₁ achieved by POD 4, frequency of PPC, postoperative length of stay (LOS), high dependency unit (HDU) LOS, sputum retention as defined by need for insertion of ‘rescue’ minitracheostomy, intensive care unit (ICU) admission and in-hospital mortality. A scoring tool to assess the frequency of PPC amenable to physiotherapy was used daily (during physiotherapy assessment). PPC was recognised by the presence of ≥4 of the following variables: chest x-ray signs of atelectasis/consolidation, elevated white cell count >11.2×10⁹/l or administration of respiratory antibiotics, temperature >38°C, positive signs of infection on sputum microbiology, oxygen saturation <90% on room air, new/changed purulent sputum production (yellow or green), physician diagnosis of pneumonia or chest infection and readmission or prolonged stay (>36 h) in ICU/HDU with problems which were respiratory in origin. This scoring tool was deemed appropriate for use in this study following comparison with other published scores.

Outcomes of subgroups defined by risk were also observed. Subjects demonstrating ≥2 independent factors predictive of the development of PPC were considered ‘high risk’. These factors included age ≥75 years, American Society of Anaesthesiologists (ASA) score ≥3, chronic obstructive pulmonary disease (COPD), smoking and body mass index (BMI) ≥30.
Statistical analysis
The sample size calculation of 37 subjects per group was based on a type I error of 5%, power of 80% and on finding a difference of 10% in mean percentage drop in FEV₁. The calculation was based on the unpublished pilot study data (n=18) which revealed a mean±SD percentage drop in FEV₁ on POD 4 of 37 ±15.24%.

Analysis of data was performed using SPSS V.17. Normally distributed continuous variables are expressed as mean±SD, skewed continuous variables as median (IQR) and categorical variables as percentages. Differences were tested with the χ² test, Fisher exact test, independent samples t test and the Mann–Whitney U test. A p value <0.05 was considered significant. The 95% CIs of differences are cited as treatment minus control. As significant differences were detected between the control and intervention groups for age and ASA score, and subsequently a significant difference in LOS was found on testing with the Mann–Whitney U test, an analysis of covariance (ANCOVA) was performed on the data to retest for significance once the observed differences had been taken into account.

Multivariate analyses were performed using both forward stepwise multiple and logistic regression. Independent variables considered potentially to have an association with outcome were gender, RCT grouping, lung cancer status, age, BMI, ASA score, COPD, smoking status, predicted FEV₁, ppo FEV₁, FEV₁/forced vital capacity (FVC) ratio and subjective preoperative activity score. Those with a p value <0.05 on univariate analysis (differences tested for as described above) were considered eligible for entry to the regression models. All continuous variables were dichotomised prior to this testing.

RESULTS
Flow of patients
The flow of patients into the study is shown in figure 1. A total of 184 patients were randomised, however three were excluded on POD 1 as the end point of PPC had been reached, and a further patient was excluded on POD 2 as this subject developed multiorgan failure requiring ICU admission; 88 patients were allocated to the control group and 92 to the intervention group.

Demographics and potential confounders
Control and intervention group demographic and risk factors are compared in table 1; significant differences were found in the age and ASA scores (p<0.05). No differences were observed in surgical procedure (p=0.987), number of unobstructed segments removed (p=0.744) or type of analgesia (p=0.277) (table 2). There was no difference observed in the modes of postoperative perceived pain scores from POD 1 to POD 4 (p=0.729, p=0.108, p=0.764, p=0.195, respectively), or in the number of patients mobilising <10 m during physiotherapy on POD 1 (p=0.303).

Effect of intervention
Only 32 of the total 180 subjects did not complete postoperative spirometry (see figure 1) due to intercostal chest drain air leak (n=16), hospital discharge (n=5) or ability to engage in test

Figure 1 Flow of patients into and through the study. PPC, postoperative pulmonary complication; VATS- video-assisted thoracic surgery.
Table 1  Control and intervention group demographic and risk factors

<table>
<thead>
<tr>
<th>Demographics, risk factors and postoperative pain</th>
<th>Control (n=88)</th>
<th>Intervention (n=92)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>47% (41)</td>
<td>49% (45)</td>
<td>0.871</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>91% (80)</td>
<td>94% (85)</td>
<td>0.536</td>
</tr>
<tr>
<td>% predicted FEV1, median (IQR)</td>
<td>75 (24)</td>
<td>77 (23)</td>
<td>0.818</td>
</tr>
<tr>
<td>Age (years), median (IQR)</td>
<td>70 (9)</td>
<td>65 (14)</td>
<td>0.003</td>
</tr>
<tr>
<td>BMI, mean±SD</td>
<td>27±5</td>
<td>27±5</td>
<td>0.966</td>
</tr>
<tr>
<td>ASA &gt;3</td>
<td>70% (61)</td>
<td>54% (50)</td>
<td>0.044</td>
</tr>
<tr>
<td>Current smoking/ ex-smokers of up to 6 weeks</td>
<td>22% (19)</td>
<td>26% (24)</td>
<td>0.595</td>
</tr>
<tr>
<td>Preoperative activity level ≥2 km</td>
<td>26% (23)</td>
<td>25% (23)</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data shown as % (n).</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASA, American Society of Anaesthesiologists; BMI, body mass index; COPD, chronic obstructive pulmonary disease.</td>
</tr>
</tbody>
</table>

(n=11). One control group subject was admitted to the ICU on POD 3 with respiratory failure. This patient subsequently died and represents the only case of in-hospital mortality. No significant difference was observed in perceived pain between the control and intervention groups during spirometry (p=0.354).

No benefit of incentive spirometry was observed, with an almost identical mean drop in FEV1 on POD 4 of 1% (p=0.817): intervention 40±16%, control 41±14% (95% CI −5.3% to 4.2%). The groups also demonstrated identical postoperative recovery of lung function with a mean difference of 1% in percentage ppo FEV1 achieved by POD 4 (p=0.744): intervention 72±19%, control 71±21% (95% CI −5.4% to 7.5%).

The frequency of PPC was 16% (n=29), and timing of when PPC was first identified is shown in figure 2. Subjects in whom PPC was recognised on POD 1 (n=5) were excluded as the end point had been met, although POD 4 spirometry was obtained where possible.

Frequency of PPC, as recognised from day 2 onwards, was 14% (24/175), with a non-significant (p=0.803) difference of 2.5% (95% CI 7.9% to 12.9%). The frequency of ‘rescue’ minitracheostomy devices was 9% (7/80) vs 3% (2/78) (18 subjects already had prophylactic minitracheostomy devices) (p=0.167). There was a significantly lower intervention group median (IQR) postoperative LOS of 6 (3) days vs 5 (3) days (p=0.047), but no significant difference in HDU LOS with a median (IQR) of 2 (2) days vs 2 (1) days (p=0.205), or ICU admission (0% vs 1%; p=0.047). As the subjects in the control group were significantly older and demonstrated higher ASA scores, ANCOVA testing was performed on LOS data to test for significance once the higher age and ASA of the control group had been taken into account. Once these factors had been adjusted for, the difference in LOS was no longer significant (p=0.186).

Subgroup analysis

There were 91 low-risk and 89 high-risk subjects. The subgroups were generally comparable, however the high-risk control subjects demonstrated a larger proportion of ASA scores ≥3 (98% vs 82%, p=0.041) and the low-risk control subjects were older (70 years vs 62 years, p=0.004).

Effect of intervention in risk subgroups

A total of 37 subjects in each subgroup completed spirometry on POD 4. There was no significant difference in perceived pain during spirometry for either the high-risk or low-risk subgroups (p=0.740 and p=0.304, respectively). In high-risk patients there was a mean between-group difference of 1% in FEV1 on POD 4 (p=0.872): intervention (n=45) 39±13%, control (n=44) 39±14% (95% CI −6.9% to 5.8%), and in the low-risk subgroup there was a mean difference of 1% (p=0.868): intervention (n=47) 42±18%, control (n=44) 43±13% (95% CI −7.7% to 6.5%). In the high-risk subgroup the mean between-group difference in percentage ppo FEV1 achieved by POD 4 was 2% (p=0.649): intervention 73±13%, control 71±22% (95% CI −5.4% to 7.3%), and in the low-risk subgroup there was no difference (p=0.970): intervention 71±24%, control 71±19% (95% CI −9.7% to 10.1%). There were no significant differences in secondary outcome measures (table 3).

Each independent risk factor for PPC may have a varying impact, therefore the outcomes for patients characterised by...
each of the five factors (age ≥75 years, ASA score ≥3, COPD, smoking and BMI ≥30) were also calculated. There were no significant differences in lung function, although there was a larger difference of 7% in favour of intervention in ppo FEV1 on POD 4 in patients with COPD (71% vs 78%). There was also a larger difference in frequency of PPC of 19% (95% CI =6.8% to 4.2%) in favour of the intervention (36% vs 17%) in patients with a history of COPD, and 15% (95% CI =-11.4% to 3.9%) in current smokers or ex-smokers of ≤6 weeks (32% vs 17%).

Multivariate analysis
Forward stepwise multiple regressions were performed to identify independent factors associated with mean drop in FEV1 and the percentage ppo FEV1 achieved on POD 4. The independent variables eligible to be entered into the regression model (p<0.05) for the dependent variable percentage drop in FEV1 were COPD, smoking status (current or ex-smoker of ≤6 weeks), predicted FEV1 ≤70%, ppo FEV1 ≤40% and FEV1/FVC obstruction ratio ≤70%. Variables eligible to be entered into the model for the dependent variable percentage of ppo FEV1 were smoking status, BMI ≥30, predicted FEV1 ≤70%, ppo FEV1 ≤40% and FEV1/FVC obstruction ratio ≤70%.

There was an acceptable fit of data (less than a total of 5% of standardised residuals were <-2 SDs or >2 SDs) and Pearson correlation coefficient r values indicated no probable collinearity. The analysis of variance F ratios were 11.231 and 15.185 (p=0.001 and p<0.001, respectively), demonstrating that the models significantly improved the ability to predict the outcome variable.

A significant contribution to the model for percentage drop was made by predicted FEV1 (p=0.001) and to the model for percentage ppo FEV1 achieved on POD 4 by ppo FEV1. The B values (~8.35 and 17.81) demonstrated that, when preoperative predicted FEV1 was >70%, the drop in FEV1 on POD 4 was less by 8%, and when ppo FEV1 was <40%, the amount of the ppo FEV1 achieved increased by 18% to 87%.

Logistic regression was performed for the dependent variable development of PPC.

Independent variables eligible to be entered into this model (p<0.05) were COPD, ppo FEV1 ≤40%, FEV1/FVC obstruction ratio ≤70%, ASA ≥3, BMI ≥30 and preoperative activity ≤400 m.

The model correctly classified 86% of patients with PPC. A significant contribution to the model was made by COPD (p=0.030). There was a positive relationship between PPC recognition and COPD, and the positive OR of 2.89 indicated that COPD was a predictor for development of PPC. The 95% CIs did not cross 1 (1.11 to 7.75), indicating that as the predictor variable increased in value, so did the odds of developing PPC.

DISCUSSION
This study is the largest fully powered RCT to investigate physiotherapy following thoracic surgery, and does not support the hypothesis that incentive spirometry either improves recovery of lung function or reduces frequency of PPC. The participants were generally representative of patients undergoing lung resection in other recent studies. Lack of improvement in recovery of lung function and effect on frequency of PPC following treatment with incentive spirometry is in agreement with the findings of two previous small studies.

The study was powered to detect a difference in postoperative loss of lung function, a smaller mean drop in lung function on POD 4 indicating enhanced recovery of lung function. However, lung function did not drop equally in all patients, as seen in the results of the multiple regression, and, interestingly, those with poorer preoperative lung function lost relatively less than those demonstrating good preoperative lung function (>70% predicted). The trend for postoperative lung volumes to improve in patients with poorer preoperative lung function may reflect the findings of other studies noting a ‘lung volume reduction effect’.

The observed loss in lung function in this study may therefore have been somewhat diluted, making possible differences more difficult to detect as they were smaller than anticipated. Patients with good preoperative lung function appear relatively ‘worse’ postoperatively; however, their outcome most likely represents the normal reduction in function seen following resection of lung tissue. An outcome such as frequency of PPC is probably a more appropriate primary outcome measure for a study such as this, and has been shown to be associated with significant clinical and economic impact.

The overall frequency of PPC in this study (16%) is comparable to that recorded in large observational studies. A non-significant difference was observed in frequency of PPC, with 95% CIs indicating no evidence of benefit with either regimen. Logistic regression demonstrated a positive relationship between the development of PPC and COPD, as might be expected, despite those with poor preoperative lung function demonstrating relatively improved postoperative lung volumes. Postoperative atelectasis and sputum retention may be worse in those with COPD due to loss of elasticity and increased preoperative secretion volume, predisposing to PPC, despite the more positive change in postoperative lung volumes often observed.

There was also no difference in recovery of lung function in the high-risk subgroups; however, there was an interesting (although non-significant) difference in frequency of PPC (9%), with the 95% CI of the difference indicating possible benefit of intervention. Further research is needed to confirm this possible benefit. Patients characterised by each independent PPC risk factor also demonstrated no significant differences in lung function, although actual differences were greater in patients with a

Table 3 Secondary outcomes in high-risk and low-risk subgroups

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>High-risk subgroup</th>
<th></th>
<th>Low-risk subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control (n=43)</td>
<td>Intervention (n=43)</td>
<td>p Value</td>
</tr>
<tr>
<td></td>
<td>Control (n=44)</td>
<td>Intervention (n=45)</td>
<td>p Value</td>
</tr>
<tr>
<td>PPC frequency</td>
<td>23% (n=10)</td>
<td>14% (n=6)</td>
<td>0.406</td>
</tr>
<tr>
<td>Insertion of ‘rescue’ minitracheostomy</td>
<td>11% (4/38)</td>
<td>0% (0/36)</td>
<td>0.115</td>
</tr>
<tr>
<td>Postoperative LOS, median (IQR)</td>
<td>6 (3)</td>
<td>5 (1)</td>
<td>0.075</td>
</tr>
<tr>
<td>HDU LOS, median (IQR)</td>
<td>2 (3)</td>
<td>2 (2)</td>
<td>0.893</td>
</tr>
<tr>
<td>ICU admission</td>
<td>0% (0)</td>
<td>2.3% (1)</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>7% (n=3)</td>
<td>11% (n=5)</td>
<td>0.714</td>
</tr>
<tr>
<td></td>
<td>7% (3/42)</td>
<td>5% (2/42)</td>
<td>1.000</td>
</tr>
</tbody>
</table>

HDU, high-dependency unit; ICU, intensive care unit; LOS, length of stay; PPC, postoperative pulmonary complication.
history of COPD. A previous study has confirmed enhanced recovery of ppo FEV₁ in patients with COPD following therapy with incentive spirometry and inspiratory muscle training.²⁷ The difference in frequency of PPC, although non-significant, was relatively large in patients with COPD at 19% and current smokers/ex-smokers of ≤6 weeks at 15%; 95% CIs indicated benefit of intervention. Significantly improved outcomes have previously been demonstrated following respiratory physiotherapy in high-risk postoperative patients with poor preoperative lung function and higher BMI.²⁸–³⁰

This study was limited by participants’ awareness of the interventions being tested; participants were not blinded, introducing a possible source of bias. This problem, however, is difficult to avoid when equipment is being tested or when a patient is performing an exercise. Also, in order specifically to compare incentive spirometry with deep breathing exercises, the regimens tested were very similar meaning that possible differences would be small. Whether or not the participants actually used the incentive spirometry device as required in terms of frequency and duration between supervised sessions is unknown. Similarly, the compliance of patients performing thoracic expansion exercises is also unknown. It is therefore possible that the results do not represent the outcomes of patients who performed interventions as desired. A compliance diary to monitor independent performance or a POD 4 survey reflecting patient usage may have been useful additions to this study.

A further limitation of this study is that the groups were not homogenous; the control group was significantly older with higher ASA scores. Larger sample size usually reduces the performance or a POD 4 survey regarding physiotherapy need to be developed further. It could be influenced of preoperative risk stratiﬁcation. Does physiotherapy reduce the incidence of postoperative complications in patients following pulmonary resection via thoracotomy? A protocol for a randomised controlled trial. J Cardiothorac Surg 2008;3:48.

CONCLUSIONS
No significant difference in postoperative lung function, frequency of PPC or LOS was observed in patients performing a physiotherapy regimen including incentive spirometry. For those at higher risk, in particular those with COPD or current/ recent ex-smokers, there were larger (although non-significant) differences in outcome which are interesting and indicate that investigations regarding physiotherapy need to be developed further. It is difficult to rule out the beneﬁt of incentive spirometry in patients at higher risk at present; however, widespread routine use of this therapy is not recommended.

Competing interests None.

Ethics approval Ethical approval for this study was granted in 2008 by the South Birmingham local research ethics committee (REC number 08/H1207/79) following approval from the local clinical governance department.

Provenance and peer review Not commissioned; internally peer reviewed.

REFERENCES