Predicting Lung Volume Reduction after Endobronchial Valve Therapy Is Maximized Using a Combination of Diagnostic Tools

T. David Koster a, Eva M. van Rikxoort b, Ralf-Harto Huebner c, Felix Doellinger d, Karin Klooster a, Jean-Paul Charbonnier b, Sri Radhakrishnan f, Felix J.F. Herth e, Dirk-Jan Slebos a

a Department of Pulmonary Diseases, University of Groningen, University Medical Center Groningen, Groningen, and b Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands; c Department of Internal Medicine/Infectious Diseases and Respiratory Medicine, and d Institute of Radiology, Charité – Universitätsmedizin Berlin, Berlin, and e Department of Pneumology and Critical Care Medicine, Thoraxklinik, Translational Lung Research Center Heidelberg, University of Heidelberg, Heidelberg, Germany; f Pulmonx Corporation, Redwood City, Calif., USA

Key Words
Bronchoscopic lung volume reduction · Endoscopic lung volume reduction · High-resolution computed tomography · One-way valve · Bronchoscopy · Collateral respiration · Chronic obstructive pulmonary disease

Abstract
Background: Bronchoscopic lung volume reduction using one-way endobronchial valves (EBVs) has been proven to be effective in patients with severe emphysema. However, the selection of patients without collateral ventilation prior to treatment is critical for procedural success. Collateral ventilation can be assessed directly with the Chartis system or indirectly using computed tomography (CT) fissure analysis.

Objectives: We retrospectively evaluated the diagnostic value of a combination of the quantitative CT interlobar fissure completeness score (FCS) and Chartis in predicting responders to EBV therapy. Methods: CT data from four prospective studies were pooled and analyzed using semiautomated software to quantify the completeness of interlobar fissures. These FCSs were compared to a reference standard of achieving ≥350 ml of target lobe volume reduction after EBV treatment. Using a receiver operating characteristic curve, optimal thresholds predictive of complete fissures (responders) and incomplete fissures (non-responders) were determined. A subgroup of patients with partially complete fissures was identified, where software had lower accuracy. The complementary value of Chartis was investigated in this group.

Results: A fissure was defined as complete (FCS >95%), incomplete (FCS <80%), or partially complete (80% < FCS < 95%). The positive predictive value (PPV) of complete fissures is 88.1%, and the negative predictive value (NPV) is 92.9%, with an overall accuracy of 89.2%. Chartis was utilized in patients with partially complete fissures, with a PPV of 82.3%, an NPV of 84.6%, and an accuracy of 83.3%.

Conclusion: Combining diagnostic tools could reduce the burden on patients and the healthcare system while providing clinicians with a better means for patient selection for EBV therapy.

© 2016 S. Karger AG, Basel
Introduction

Approximately 8% of the global population suffers from chronic obstructive pulmonary disease (COPD) [1]. Severe emphysema is a subpopulation of highly symptomatic COPD patients suffering from dyspnea due to lung hyperinflation, with a reduced life expectancy and with very few effective therapeutic alternatives [1]. Bronchoscopic lung volume reduction using endobronchial valves (EBVs) has recently been shown to deliver substantial, clinically relevant improvements to emphysema patients with hyperinflation. In this therapy, one-way EBVs are delivered via a bronchoscope, with the goal of occluding a lobe and collapsing hyperinflated lung regions by allowing trapped gas to exit while preventing inflow during inhalation [2]. The key predictor for response to EBV therapy has been shown to be the selection of patients, with low collateral ventilation between the target lobe to be occluded and the adjacent lobe(s) [3, 4].

There are several methods to measure the presence of collateral ventilation [5]. A well-studied direct measurement of collateral airflow is the Chartis system® (Pulmonx, Redwood City, Calif., USA), which is an endobronchial pulmonary assessment system [6, 7]. However, performing this measurement requires a bronchoscopy, can include procedural challenges, and, when used in a broader population, will result in a substantial number of patients undergoing a bronchoscopy who will not receive EBVs [6–9]. Alternately, an indirect assessment of collateral ventilation is anatomical fissure analysis using computed tomography (CT). A complete fissure on CT is considered a surrogate for the absence of collateral ventilation. Goldin and Abtin [10] proposed that, if >90% of the fissure could be visualized on at least one axis (sagittal, axial, or coronal view), the fissure should be considered complete, implying the absence of significant collateral ventilation between the adjoining lobes. Several studies have adopted this, without subsequent refinement of this ad hoc threshold [3, 8, 11, 12]. However, outside the experienced radiology core lab setting, visual estimation of fissure completeness has been associated with a rather poor interobserver agreement [13].

Several computerized semiautomated methods to quantify the completeness of fissures have been developed, which claim to reduce the interobserver variability or complement the visual read of an experienced radiologist [14, 15]. These methods are more efficient and have an improved reliability over visual methods. A recent study concluded that the optimal treatment algorithm might be a combination approach, using radiology for an anatomical classification based on fissure completeness, and a physiologic measure of airflow in patients with partially complete fissures [16].

In this study, we evaluated the diagnostic value of the combination of a quantitative CT (QCT) interlobar fissure completeness score (FCS) and Chartis in predicting responders to EBV therapy, with the aim to develop a diagnostic workflow providing clinicians with a better tool for patient selection.

Methods

Subjects and Study Design

This study is a retrospective analysis of 547 patients based on pooled data collected from four prospective studies. The clinical trial registrations and procedures of these studies have been described in detail before [3, 4, 6, 12, 17]. One dataset is derived from a postmarketing registration study with the goal to evaluate the outcome after endoscopic lung volume reduction, which has not been previously published but has been approved by the Ethics Committee of Charité University Hospital, Berlin, Germany (EA1/136/13). A part of this German dataset has been used to retrospectively assess the occurrence of pneumothorax after EBV treatment [18].

Prior to the analysis, all patient records and information were deidentified and rendered anonymous, and the analysts performing the QCT evaluation were provided baseline scans only and were blinded to the posttreatment clinical outcomes. Cases were deemed eligible if they received EBV therapy and if follow-up CT scans were performed. To reflect routine clinical practice, CT scans from 16 sites in the United States and 17 European sites collected from 2004 through 2015 were considered for analysis. Cases were excluded if there was a lack of procedural success placing valves, as confirmed by bronchoscopy or an independent radiology core lab, since this is a confounding variable for the predictive ability of any diagnostic tool used. Cases were also excluded if the quality of the scans were insufficient for QCT analysis, for example thick slices or missing slices, resulting in an incomplete reconstruction.

Image Analysis

Since the goal of QCT used in conjunction with EBV therapy is lung volume reduction, target lobe volume reduction (TLVR) was considered the gold standard for determining the accuracy of predicting response. A TLVR of ≥350 ml has previously been established as a measure of procedural success [6]. QCT analysis was performed on all baseline CT scans at Thirona (Thirona BV, Nijmegen, The Netherlands) using Thirona LungQ version 1.0.0 to assess fissure completeness and lobar tissue destruction at baseline for each subject. In each scan, the lungs, pulmonary fissures, and pulmonary lobes were automatically segmented, visually checked, and edited by trained medical analysts [19–22]. Each scan was read by two medical analysts, with the first analyst editing the segmentation results where needed and the second analyst checking the results. Based on the results of the lobe and fissure segmentations, an FCS (fig. 1) was computed for each lobe as the percentage of the lobar boundaries defined by a fissure [15]. Within the segmented

Predicting Lung Volume Reduction after EBV Therapy

Respiration 2016;92:150–157
DOI: 10.1159/000448849
lobes, attenuation thresholding (in Hounsfield units, HU) was performed to quantify the emphysema severity as the percentage of voxels below –910 HU, as previously defined [15].

Once the FCSs were finalized, precalculated TLVR scores, determined from the change in volume of the target lobe before and after EBV therapy (fig. 2), were utilized to dichotomize all subjects into responders and non-responders, with a non-responder defined as a subject with <350 ml reduction and a responder defined as a subject with ≥350 ml volume reduction.

Statistical Analysis

The FCS was evaluated for its ability to predict a responder (TLVR ≥350 ml), using a receiver operating characteristics (ROC) curve. Sensitivity and specificity were calculated for specific FCSs in 5% increments to identify the optimal threshold which maximized the sensitivity and specificity. The positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, and accuracy of the QCT system were computed by applying the optimal FCS threshold.

To investigate the combined use of QCT and Chartis, a lower FCS threshold was determined by optimizing specificity at high sensitivity, minimizing false negatives below this threshold.

The application of both FCS thresholds results in three groups: complete fissures (FCS > optimal threshold), incomplete fissures (FCS < lower threshold), and partially complete fissures (optimal threshold < FCS < lower threshold). In the partially complete fissure group, where sensitivity and specificity of QCT are suboptimal, the use of Chartis for decision to treat is analyzed.

Results

Valve treatment was initially performed in 547 subjects, and, as shown in figure 3, we eventually included 217 subjects in our analysis. The baseline characteristics of the included subjects are presented in table 1.
In the 217 patients, the right lower lobe was targeted in 7% of the cases, the right upper lobe in 33%, the left upper lobe in 28%, the left lower lobe in 27%, the middle lobe in 5%, and the right upper or lower lobe together with the middle lobe in 4.6 and 0.5%, respectively.

The ROC curve showed an area under the curve of 0.86. The best cutoff to maximize sensitivity and specificity was an FCS of 95% (fig. 4).

The PPV, NPV, sensitivity, specificity, and accuracy of the QCT system using the FCS >95% were computed and are provided in table 2. Based on the ROC curve (fig. 4), the lower threshold representing incomplete fissures was determined to be FCS <80%. Applying a <80% do-not-treat and a >95% treat algorithm to the entire dataset, the algorithm had an overall accuracy of 89.2%, not including the 80% < FCS < 95% population (table 2).

To explore the possibility of combining diagnostic tools, all subjects with partially complete fissures (i.e. 80% < FCS < 95%) were further evaluated for the presence of a baseline Chartis assessment, and 30 subjects, all from one study, had evaluable Chartis assessments [6]. Applying previously described waveform analysis criteria for predicting responders and applying the same criteria as above (≥350 ml TLVR) for defining a responder [18, 23], Chartis was found to have an accuracy of 83.3% in this subgroup of partially complete fissures, as shown in table 2.

**Discussion**

**Interpretation**

The treatment of severe emphysema with EBV therapy has been proven effective in randomized controlled trials [4, 11]. However, optimal patient selection continues to be an area of clinical research. QCT analysis is a rapidly emerging field in pulmonary diagnosis and a potential...
powerful tool in emphysema disease staging for advanced therapies.

In this study, we identified thresholds for applying a proprietary QCT algorithm for classifying patients, and a strategy for selectively using Chartis for increased diagnostic yield. In patients with complete fissures (>95% FCS), the algorithm provides an 88.1% PPV; therefore, there may be little diagnostic value in subjecting these patients to an additional physiological Chartis assessment, and these patients may be treated with EBV without further evaluation. Since for patients with incomplete fissures (<80% FCS) the algorithm provides a 92.9% NPV, these patients can be excluded from further EBV treatment or additional invasive diagnostics without denying treatment to a large group of patients. Thus, QCT can be used as a prescreening tool for all patients, and Chartis can be used selectively, only on patients with partially complete fissures.

**Comparison to the Literature**

A few studies have investigated the diagnostic value of fissure completeness for prescreening prior to EBV therapy. Gompelmann et al. [12] visually reviewed CT scans for evaluating fissure completeness above or below a 90% threshold. A similar reference standard of volume reduction ≥350 ml was applied to define responders. They found an accuracy of 77%, a sensitivity of 75%, and a specificity of 78.8%, which were slightly lower than those in our experience. This could be due to multiple factors – the fissure threshold applied or the variability in visually assessing the fissures, though these scans were reviewed at a radiology core lab. This study may indicate the issues with nonexperienced readers in the clinical setting, estimating the completeness of the fissure while manually scrolling through the 300–500 slices of a high-resolution CT. Schumann et al. [8] used quantitative imaging software to retrospectively compare its accuracy to the Chartis system in 134 patients using the same reference standard of TLVR ≥350 ml. Based on regression analysis of 34 QCT variables, three predictors for successful lung volume reduction were identified: fissure completeness, low attenuation clusters, and a vascular index. Using fissure completeness as a single feature yielded an area under the ROC curve of 0.75, which was not statistically different from using all three features (area under the ROC curve of 0.80). Applying a bimodal distribution of patients above and below a single threshold of 90% fissure completeness, the overall accuracy was 75.8%, with a sensitivity of 83.3% and a specificity of 66.7%, which was similar to the results of QCT alone in our study. The third study was performed by de Oliveira et al. [16], who retrospectively analyzed 38 treated patients. They investigated the relationship between fissure completeness and clinically relevant lobar volume reduction (≥350 ml), using a QCT analysis. A fissure completeness >90% was found to have a PPV of 90.5%, whereas <75% fissures had a NPV of 100%. The accuracy of the QCT in the middle zone (between 75 and 90% fissure completeness) was lower, and the authors suggested the use of the Chartis system in this middle zone.

**Table 2. Predictive values of FCS and Chartis**

<table>
<thead>
<tr>
<th>FCS threshold</th>
<th>Responder</th>
<th>Non-responder</th>
<th>Prevalence</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCS &gt;95%</td>
<td>119</td>
<td>16</td>
<td>135</td>
<td>88.1%</td>
<td></td>
</tr>
<tr>
<td>FCS &lt;95%</td>
<td>26</td>
<td>56</td>
<td>82</td>
<td>68.3%</td>
<td></td>
</tr>
<tr>
<td>FCS &gt;80%</td>
<td>142</td>
<td>33</td>
<td>175</td>
<td>81.1%</td>
<td></td>
</tr>
<tr>
<td>FCS &lt;80%</td>
<td>3</td>
<td>39</td>
<td>42</td>
<td>92.9%</td>
<td></td>
</tr>
<tr>
<td>80% &lt; FCS &lt; 95%</td>
<td>3</td>
<td>16</td>
<td>135</td>
<td>88.1%</td>
<td></td>
</tr>
<tr>
<td>FCS &gt;80%</td>
<td>119</td>
<td>33</td>
<td>82</td>
<td>82.3%</td>
<td></td>
</tr>
<tr>
<td>FCS &lt;80%</td>
<td>3</td>
<td>39</td>
<td>42</td>
<td>84.6%</td>
<td></td>
</tr>
<tr>
<td>Chartis system</td>
<td>14</td>
<td>3</td>
<td>17</td>
<td>82.3%</td>
<td></td>
</tr>
<tr>
<td>CV–</td>
<td>2</td>
<td>11</td>
<td>13</td>
<td>84.6%</td>
<td></td>
</tr>
</tbody>
</table>

Se = Sensitivity; Sp = specificity; CV– = no collateral ventilation measured by Chartis; CV+ = collateral ventilation measured by Chartis.

1 Predictive values and accuracy for an FCS threshold of 95%.
2 Predictive values and accuracy for an FCS threshold of 80%.
3 Predictive values and accuracy for a 80% < FCS < 95%.
4 Predictive values and accuracy of the Chartis system in a subset of patients within this study, with valid Chartis assessments and evaluable CT scans.

Table 2 shows predictive values of FCS and Chartis for different thresholds of fissure completeness. The table includes the number of responders and non-responders, the prevalence, and the positive and negative predictive values (PPV and NPV) for each threshold. The sensitivity (Se) and specificity (Sp) are also provided for each category. The accuracy of the QCT and Chartis system is also listed for comparison.
Limitations of the Study

One of the limitations of this study is that it is a retrospective study pooling multiple trials with slightly different inclusion criteria. The more recent studies only included patients with visually complete fissures on CT to avoid diagnostic bronchoscopies with the Chartis system, whereas the earlier studies did not use fissure completeness for patient selection. Depending on the FCS threshold used, our dataset had anywhere from 62% (with FCS >95%) to 79% (with FCS >80%) patients with complete fissures, which is not representative of the overall population [5]. Using the FSC thresholds of 95% and 80% in the subgroup of 123 patients from two studies [3, 6, 17], where patients were not visually preselected based on their fissure status, we found the distribution to be 38% complete fissures, 33% incomplete fissures, and 29% partially complete fissures. This further strengthens the argument that the number of the patients with partially complete fissures in a nonselected cohort is not insignificant and should be evaluated with the Chartis to confirm the absence of collateral ventilation, prior to EBV therapy.

The entire population can thus be divided into three groups – complete fissures (treat), incomplete fissures (do not treat), and partially complete fissure (perform Chartis to confirm the absence of collateral ventilation). Applying the PPV, NPV, and accuracies from table 2 to the patient distribution described above results in a diagnostic workflow with an overall accuracy of 89.5% which is higher than Chartis or QCT alone. It results in fewer potential responders being denied treatment, and as many as 71% fewer diagnostic bronchoscopies compared to a Chartis-only screening strategy. This is visualized in figure 5, which is a hypothetical reproduction of the workflow. We used the 123 patients from two studies [3, 6, 17] and their distribution to simulate three different possible diagnostic approaches for treatment selection: Chartis only, QCT only, or the combination of Chartis and QCT. As shown in figure 5, there are some important differences: the combined approach results in higher accuracy, the Chartis approach requires all patients to undergo bronchoscopies, and using only QCT will result in significantly more potential responders being denied treatment.

Fig. 5. Hypothetical diagnostic workflow based on the study results, comparing Chartis only to QCT only and to the combination of Chartis and QCT.
The chosen reference standard for establishing the accuracy of fissure completeness is the achievement of ≥350 ml of volume reduction. This value has been used by several other studies on this subject and is considered a clinically significant and meaningful lung volume reduction compared to controls without an intervention [4, 6, 8, 12]. An alternate reference standard is surgical inspection of the fissure or a completely collapsed lobe, both of which were not realistic in this retrospective study. Moreover, since the goal of the treatment is lung volume reduction, this reference standard is clinically relevant and quantifiable. Given the various studies involved and the retrospective nature of this study, follow-up scans to calculate TLVR were acquired at various time points from 1 to 6 months. This, however, was not considered a bias, since the occurrence of significant TLVR, irrespective of the time to follow-up, is considered a successful response.

Another potential source of bias is the lack of technical success in placing valves. A misplaced valve or missed segmental airway compromises the entire treatment and is unlikely to cause lobar volume reduction, thus unfairly penalizing the accuracy of the diagnostic prediction. An example of a misplaced valve is shown in figure 6. To account for this, CT scans from four studies were checked for procedural errors by two authors (E.M.v.R. and J.-P.C.), with discordant cases being verified by T.D.K. and D.-J.S. [3, 4, 6, 17]. A total of 23 cases were found to have procedural errors, leaving 194 patients. Of this group, there were 135 responders. Removing the procedural errors leads to better predictive values of the QCT fissure analysis: the PPV with FCS >95% is 94.1%, the NPV with FCS <80% is 92.1%.

Finally, since we have applied a different QCT software algorithm and have challenged the status quo of the previously assumed 90% fissure completeness threshold with a larger dataset, our thresholds are not directly comparable to historical references, and further research is required to prospectively validate them. Nevertheless, this is the largest dataset of patients treated with EBV evaluated with QCT software, covering a broad range of CT scan acquisition protocols, geographic locations, and scanners over a long time period and incorporates data from the largest prospective trials of EBV therapy to date. Although a prospective diagnostic study is preferred, we believe that these results are relevant and robust.

In conclusion, QCT analysis can be utilized to classify severe emphysema patients into three 'fissure status' groups. This can be used to thereafter only perform a functional assessment of collateral flow with the Chartis system on those patients with partially complete fissures.

Acknowledgements

We like to thank the original US-VENT, EU-VENT, CHARTIS, and STELVIO study groups. We would also like to thank Tami Crabtree, Dominik Herzog, Dorothea Thellig, and Ryan Olivera for statistical and data analysis assistance.

Fig. 6. Example of an ideal valve placement and lack of lobar occlusion as visualized on CT. Red arrows indicate a segmental airway, where valve placement was missed in axial (a), coronal (b), and sagittal (c) view.
Predicting Lung Volume Reduction after EBV Therapy

DOI: 10.1159/000448849

References