Reproducibility of optical coherence tomography airway imaging

Miranda Kirby,[1,2] Keishi Ohtani,[3] Taylor Nickens,[1] Rosa Maria Lopez Lisbona,[4] Anthony M. D. Lee,[5] Tawimas Shaipanich,[5] Pierre Lane,[5] Calum MacAulay,[5] Stephen Lam,[5,6] and Harvey O. Coxson[1,2,*]

[1]*Centre for Heart Lung Innovation, St. Paul's Hospital, Vancouver, British Columbia, Canada*
[2]*Department of Radiology, University of British Columbia, Vancouver, British Columbia, Canada*
[3]*Department of Surgery, Tokyo Medical University, Tokyo, Japan*
[4]*Department of Respirology, Bellvitge University Hospital, l'Hospitalet de Llobregat, Barcelona, Spain*
[5]*Imaging Unit, Integrative Oncology Department, British Columbia Cancer Agency Research Centre, Vancouver, British Columbia, Canada*
[6]*Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada*
[*]Harvey.Coxson@hli.ubc.ca

Abstract: Optical coherence tomography (OCT) is a promising imaging technique to evaluate small airway remodeling. However, the short-term insertion-reinsertion reproducibility of OCT for evaluating the same bronchial pathway has yet to be established. We evaluated 74 OCT data sets from 38 current or former smokers twice within a single imaging session. Although the overall insertion-reinsertion airway wall thickness (WT) measurement coefficient of variation (CV) was moderate at 12%, much of the variability between repeat imaging was attributed to the observer; CV for repeated measurements of the same airway (intra-observer CV) was 9%. Therefore, reproducibility may be improved by introduction of automated analysis approaches suggesting that OCT has potential to be an *in-vivo* method for evaluating airway remodeling in future longitudinal and intervention studies.

OCIS codes: (170.0170) Medical optics and biotechnology; (170.4500) Optical coherence tomography.

References and links

1. "Global Initiative for Chronic Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease," (2013. Available from: *http://www.goldcopd.org/.),* p. Available from: *http://www.goldcopd.org/.*
2. J. J. Armstrong, M. S. Leigh, D. D. Sampson, J. H. Walsh, D. R. Hillman, and P. R. Eastwood, "Quantitative upper airway imaging with anatomic optical coherence tomography," Am. J. Respir. Crit. Care Med. **173**(2), 226–233 (2006).
3. L. P. Hariri, M. B. Applegate, M. M. Kenudson, E. J. Mark, B. D. Medoff, A. D. Luster, B. E. Bouma, G. J. Tearney, and M. J. Suter, "Volumetric optical frequency domain imaging of pulmonary pathology with precise correlation to histopathology," Chest **143**, 64–74 (2013).
4. M. Tsuboi, A. Hayashi, N. Ikeda, H. Honda, Y. Kato, S. Ichinose, and H. Kato, "Optical coherence tomography in the diagnosis of bronchial lesions," Lung Cancer **49**(3), 387–394 (2005).
5. H. O. Coxson, B. Quiney, D. D. Sin, L. Xing, A. M. McWilliams, J. R. Mayo, and S. Lam, "Airway wall thickness assessed using computed tomography and optical coherence tomography," Am. J. Respir. Crit. Care Med. **177**(11), 1201–1206 (2008).
6. J. P. Williamson, R. A. McLaughlin, W. J. Noffsinger, A. L. James, V. A. Baker, A. Curatolo, J. J. Armstrong, A. Regli, K. L. Shepherd, G. B. Marks, D. D. Sampson, D. R. Hillman, and P. R. Eastwood, "Elastic properties of the central airways in obstructive lung diseases measured using anatomical optical coherence tomography," Am. J. Respir. Crit. Care Med. **183**(5), 612–619 (2011).
7. J. Yin, G. Liu, J. Zhang, L. Yu, S. Mahon, D. Mukai, M. Brenner, and Z. Chen, "In vivo early detection of smoke-induced airway injury using three-dimensional swept-source optical coherence tomography," J. Biomed. Opt. **14**(6), 060503 (2009).
8. R. G. Michel, G. T. Kinasewitz, K. M. Fung, and J. I. Keddissi, "Optical coherence tomography as an adjunct to flexible bronchoscopy in the diagnosis of lung cancer: a pilot study," Chest **138**(4), 984–988 (2010).
9. J. M. Ridgway, G. Ahuja, S. Guo, J. Su, U. Mahmood, Z. Chen, and B. Wong, "Imaging of the pediatric airway using optical coherence tomography," Laryngoscope **117**(12), 2206–2212 (2007).

10. S. C. Whiteman, Y. Yang, D. Gey van Pittius, M. Stephens, J. Parmer, and M. A. Spiteri, "Optical coherence tomography: real-time imaging of bronchial airways microstructure and detection of inflammatory/neoplastic morphologic changes," Clin. Cancer Res. **12**(3), 813–818 (2006).
11. K. Wijesundara, C. Zdanski, J. Kimbell, H. Price, N. Iftimia, and A. L. Oldenburg, "Quantitative upper airway endoscopy with swept-source anatomical optical coherence tomography," Biomed. Opt. Express **5**(3), 788–799 (2014).
12. F. B. Lazarow, G. S. Ahuja, A. Chin Loy, E. Su, T. D. Nguyen, G. K. Sharma, A. Wang, J. Jing, Z. Chen, and B. J. Wong, "Intraoperative long range optical coherence tomography as a novel method of imaging the pediatric upper airway before and after adenotonsillectomy," Int. J. Pediatr. Otorhinolaryngol. **79**(1), 63–70 (2015).
13. Y. Nakano, J. C. Wong, P. A. de Jong, L. Buzatu, T. Nagao, H. O. Coxson, W. M. Elliott, J. C. Hogg, and P. D. Paré, "The prediction of small airway dimensions using computed tomography," Am. J. Respir. Crit. Care Med. **171**(2), 142–146 (2005).
14. H. U. Kauczor, M. Ebert, K. F. Kreitner, H. Nilgens, R. Surkau, W. Heil, D. Hofmann, E. W. Otten, and M. Thelen, "Imaging of the lungs using 3He MRI: preliminary clinical experience in 18 patients with and without lung disease," J. Magn. Reson. Imaging **7**(3), 538–543 (1997).
15. A. M. Lee, M. Kirby, K. Ohtani, T. Candido, R. Shalansky, C. MacAulay, J. English, R. Finley, S. Lam, H. O. Coxson, and P. Lane, "Validation of airway wall measurements by optical coherence tomography in porcine airways," PLoS One **9**(6), e100145 (2014).
16. G. Atkinson and A. M. Nevill, "Statistical Methods For Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine," Sports Med. **26**(4), 217–238 (1998).
17. J. Jekel, D. Katz, D. Wild, and J. Elmore, *Epidemiology, biostatistics, and preventive medicine* (Elsevier Inc., Philadelphia, 2007).
18. J. C. Hogg, P. T. Macklem, and W. M. Thurlbeck, "Site and nature of airway obstruction in chronic obstructive lung disease," N. Engl. J. Med. **278**(25), 1355–1360 (1968).
19. M. Cazzola, W. MacNee, F. J. Martinez, K. F. Rabe, L. G. Franciosi, P. J. Barnes, V. Brusasco, P. S. Burge, P. M. Calverley, B. R. Celli, P. W. Jones, D. A. Mahler, B. Make, M. Miravitlles, C. P. Page, P. Palange, D. Parr, M. Pistolesi, S. I. Rennard, M. P. Rutten-van Mölken, R. Stockley, S. D. Sullivan, J. A. Wedzicha, E. F. Wouters, and S. American Thoracic; American Thoracic Society; European Respiratory Society Task Force on outcomes of COPD, "Outcomes for COPD pharmacological trials: from lung function to biomarkers," Eur. Respir. J. **31**(2), 416–469 (2008).
20. R. Pellegrino, G. Viegi, V. Brusasco, R. O. Crapo, F. Burgos, R. Casaburi, A. Coates, C. P. van der Grinten, P. Gustafsson, J. Hankinson, R. Jensen, D. C. Johnson, N. MacIntyre, R. McKay, M. R. Miller, D. Navajas, O. F. Pedersen, and J. Wanger, "Interpretative strategies for lung function tests," Eur. Respir. J. **26**(5), 948–968 (2005).
21. B. R. Celli, W. MacNee, and A. E. T. Force; ATS/ERS Task Force, "Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper," Eur. Respir. J. **23**(6), 932–946 (2004).
22. A. Dirksen, J. H. Dijkman, F. Madsen, B. Stoel, D. C. Hutchison, C. S. Ulrik, L. T. Skovgaard, A. Kok-Jensen, A. Rudolphus, N. Seersholm, H. A. Vrooman, J. H. Reiber, N. C. Hansen, T. Heckscher, K. Viskum, and J. Stolk, "A randomized clinical trial of alpha(1)-antitrypsin augmentation therapy," Am. J. Respir. Crit. Care Med. **160**(5), 1468–1472 (1999).

1. Introduction

The structural changes and thickening of the airway wall components that occur in individuals with chronic respiratory diseases – known as airway remodeling – are thought to be responsible for many of the adverse outcomes associated with disease. However, our understanding of airway wall remodeling is based on studies that directly examine lung tissue, typically post-mortem, or indirectly make inferences about structure by investigating the functional effects of airway remodeling, such as spirometry measurements of airflow limitation. Even though measurements of airflow limitation are valuable and are used as markers of airway disease [1], these tests provide only global assessments of pulmonary function and no structural information at all. Therefore, to understand airway remodeling direct measurements of the airway wall in living subjects are required. Furthermore, airway diseases, such as chronic obstructive pulmonary disease (COPD) and asthma, are known to be highly regional affecting only certain airways in the lung, and, as such, global functional measurements cannot provide us with any detailed structural or regional information to better understand, characterize and possibly treat the specific remodeled airway walls.

Optical coherence tomography (OCT) is a minimally invasive imaging technique that can provide higher resolution images of small airway wall structure [2–12] not possible with other pulmonary imaging approaches, including computed tomography (CT) [13] and magnetic resonance imaging (MRI) with inhaled contrast agents [14]. Early investigations using OCT demonstrated not only that airway wall measurements correlated significantly with CT and

spirometry [5], but OCT can identify airway wall components *in vivo* that until now have only been visualized using histology [3, 4].

Despite this important early work demonstrating the potential of OCT for airway imaging in subjects with chronic respiratory diseases, there are no reports on the reproducibility of OCT for repeated airway wall measurements of the same peripheral airway paths in living humans. In patients at risk and with COPD, the ability to visualize and quantify the structural changes that occur within the small airway wall may help to inform treatment strategies to target the underlying disease mechanisms, as well as to help monitor response to specific treatments. Furthermore, indentifying individuals with significant small airway remodeling may help stratify patients for more targeted therapies in clinical trials, to ultimately obtain better therapeutic outcomes. This may be of particular importance in patients with early and mild disease, where spirometry measurements do not reflect patient symptoms or functional limitation and early intervention may slow disease progression. However, to better understand the role of OCT for investigating how airway remodeling is modified by therapy, the reproducibility of OCT must be established. The objective of this study was threefold: 1) to determine OCT inter- and intra-observer airway wall measurement reproducibility, 2) to determine the OCT probe insertion-reinsertion reproducibility for identifying and evaluating the same airway path twice within the same imaging session, and, 3) to determine OCT insertion-reinsertion airway wall measurement reproducibility. Establishing the insertion-reinsertion technique and airway wall measurement reproducibility will allow for smallest detectable difference estimates and sample size calculations to help guide future longitudinal or intervention studies.

2. Materials and methods

2.1 Study subjects

Subjects participating in an ongoing National Cancer Institute sponsored chemoprevention trial between 40 and 80yrs were enrolled in this study. All subjects were current or former smokers. COPD was defined by post-bronchodilator spirometry in accordance with the Global initiative for chronic Obstructive Lung Disease (GOLD) criteria [1]. The study was approved by the University of British Columbia's Review of Ethics Board, and written informed consent was obtained for all subjects (REB H10-00226 & H07-01393).

2.2 OCT image acquisition

OCT imaging was performed using a frequency domain swept-source OCT system (Lightlabs C7XR, St. Jude Medical, Inc., St. Paul, MN, USA) and a 0.9mm diameter optical catheter (C7 Dragonfly Imaging Catheter, St. Jude Medical Inc.) enclosed in a 1.5mm diameter clear Pebax sheath [15]. The OCT catheter was inserted through the biopsy channel of the bronchoscope into a sub-segmental airway in the lower, middle or upper lobe. Once at the sub-segmental bronchus entrance, the probe was gently advanced until the internal diameter of the airway was equal to the 1.5mm outer diameter of the probe. Three-dimensional imaging of a 5cm airway path distal to the entrance of the sub-segmental bronchus was obtained using a computer controlled pull-back of the rotating probe at 1cm/sec. The catheter was removed and the bronchoscope was withdrawn to the trachea. The bronchoscope was then repositioned at the same sub-segmental bronchus entrance and the catheter was re-inserted into the same airway for repeat imaging.

2.3 OCT image analysis

To determine if the OCT probe was positioned in the exact same airway on the second imaging procedure, the three-dimensional airway paths were compared and a complete match was defined to have occurred when all branch-points were matched in both proximal and the peripheral sections of the airway path (Fig. 1). A partial match was defined to have occurred when branch-points were matched in the proximal part of the airway segment, but not the peripheral section (Fig. 2).

Each airway segment was classified according the standard bronchial numbering system and grouped according to lobe. Airway wall dimensions were assessed using three consecutive OCT frames in each airway segment. The lumen area (Ai) and outer wall area (Ao) were manually segmented using ImageJ software (National Institutes of Health, Bethesda, MD) and the mean Ai and Ao were used to calculate the mean airway wall area (WA = Ai-Ao). Wall thickness (WT) was also calculated by the subtraction of the average lumen diameter from the average outer wall diameter; average lumen diameter was calculated according to Eq. (1):

$$\text{Average lumen diameter} = 2\sqrt{\frac{Ai}{\pi}} \tag{1}$$

and average outer wall diameter was calculated according to according to Eq. (2):

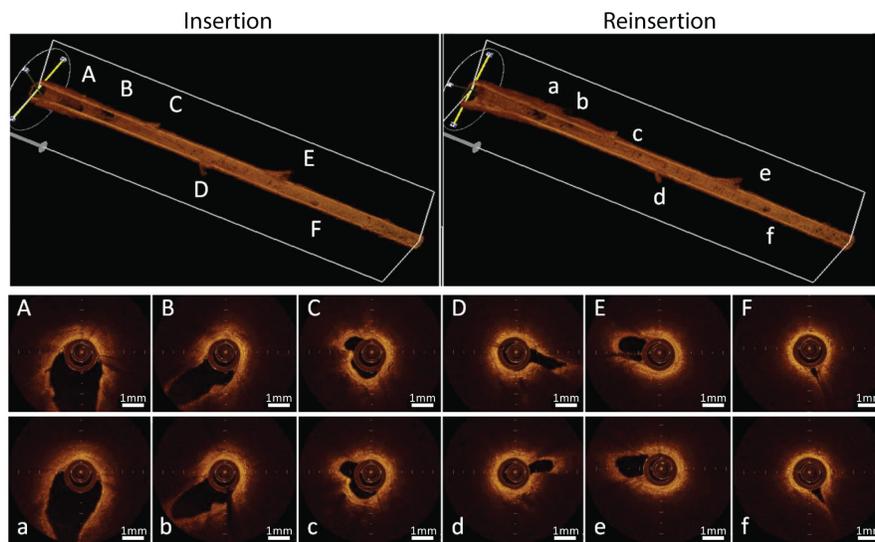$$\text{Average outer wall diameter} = 2\sqrt{\frac{Ao}{\pi}} \tag{2}$$

Fig. 1. OCT imaging of a completely matched airway. Three-dimensional optical coherence tomography (OCT) images of a sub-segmental airway path obtained by the first insertion (A) and second insertion (B) of the OCT catheter. The branch points are shown by capital and lower case letters respectively. The lower images are cross-sectional OCT images of each branch point.

OCT airway segmentation was performed by three observers. One observer was a pulmonologist with experience performing the OCT procedure and quantitative analysis; the second observer was an expert in quantitative imaging with experience performing the quantitative analysis; the third observer had no previous experience with the technique, but was provided with brief training. Observer 3 performed three rounds of segmentation, blinded to all subject demographic information and imaging time-points, with a minimum of 24 hours between repeated segmentation rounds to minimize memory bias.

2.4 Statistical analysis

Statistical comparison of demographic and spirometry measurements between subjects with and without COPD were performed using Mann Whitney t-tests for continuous variables (GraphPad Software Inc, San Diego, CA, USA). A Fisher's exact test was used for all statistical comparisons of categorical variables. Inter-observer, intra-observer and insertion-reinsertion measurement reproducibility were determined using Pearson correlation

coefficients (r) and Bland-Altman analysis (GraphPad Software Inc). Two-way mixed-effects repeated measures analysis of variance was performed to determine if there were significant differences between measurements made by the three observers, and the insertion-reinsertion measurements using SAS 9.2 software (SAS Institute, Cary, NC).
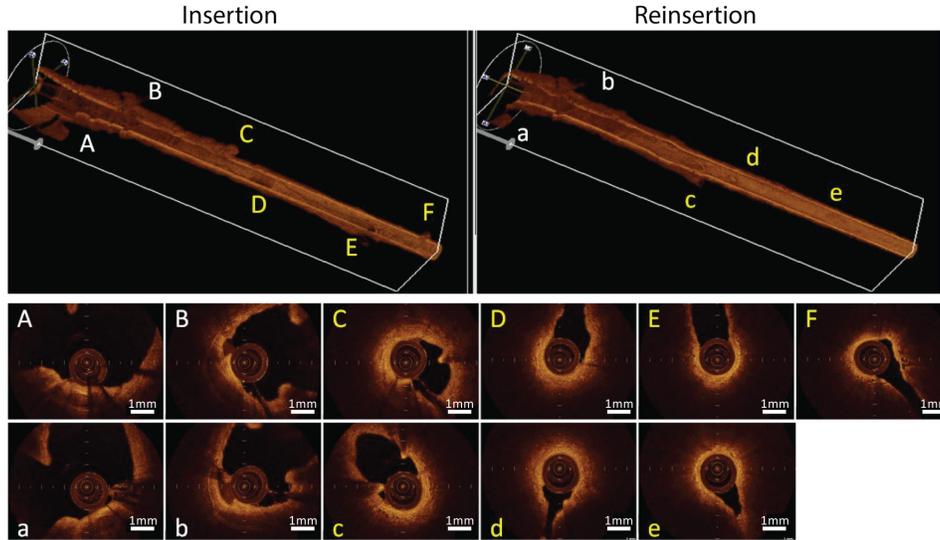

Fig. 2. OCT imaging of a partially matched airway. Three-dimensional optical coherence tomography (OCT) images of sub-segmental airway path obtained by the first insertion (A) and second insertion (B) of the OCT catheter. Branch points A and B are identical but the catheter takes a different path at branch point C and, therefore, all peripheral branch points are different (yellow letters).

For the single observer that performed repeated measurements, the intra-observer and insertion-reinsertion measurement intraclass correlation coefficient (ICC) (MedCalc Software, Ostend, Belgium) and coefficient of variation (CV) were also generated. Using the insertion-reinsertion measurement ICC as previously described [16], the smallest detectable difference (SDD), defined as the magnitude of change required to exceed the technique and measurement error for repeat imaging at two different time-points with 95% confidence ($\alpha = 0.05$, $Z_\alpha = 1.96$), was calculated as shown in Eq. (3):

$$SDD = 1.96\sqrt{2}SD_{\text{Insertion}}\sqrt{1 - ICC_{\text{Insertion-Reinsertion}}} \qquad (3)$$

where $SD_{\text{Insertion}}$ is the standard deviation of WT measurements at the first time-point and $ICC_{\text{Insertion-Reinsertion}}$ is the intraclass correlation coefficient of insertion-reinsertion WT measurements.

A sample size calculation [17] was performed to help guide future controlled treatment studies using the variation in the insertion-reinsertion WT measurements for subjects at risk and with COPD. Since no interventions were performed in this study, we determined the number of subjects (n) that would be required in a controlled trial to detect significant differences (δ) for a range of effect sizes typically measured in clinical trials following an intervention for WT measurements between baseline and follow-up in a control group (subjects at risk and with COPD receiving no treatment) and a treatment group with 95% confidence ($\alpha = 0.05$, $Z_\alpha = 1.96$) and 80% power ($\beta = 0.20$, $Z_\beta = 0.84$) according to Eq. (4):

$$n = \frac{2(Z_\alpha + Z_\beta)^2 SD_{Diff}^{\ 2}}{\delta^2} \qquad (4)$$

where SD_{Diff} was the standard deviation of the difference between baseline and follow-up in the no treatment group; SD_{Diff} was estimated using the standard deviation of the difference between the average of the three observer's insertion-reinsertion WT measurements; δ was the difference between the baseline and follow-up mean change for the control and treatment group. All results were considered statistically significant when the probability of making a Type I error was less than 5% ($p < 0.05$).

3. Results

The subject demographic information for 38 current or ex-smokers are shown in Table 1. There were no significant differences between current or ex-smokers with and without COPD with respect to age ($p = 0.27$), smoking status ($p = 0.73$), and pack-years ($p = 0.26$), however there were significantly fewer females with COPD ($p = 0.01$).

Table 1. Subject Demographics

Parameter (± SD)	All n = 38	Control n = 26	COPD n = 12	Significance of Difference (p)*
Age, yr (± SD)	62 (9)	61 (8)	65 (9)	0.27
Female, n (%)	16 (42)	7 (27)	9 (75)	0.01
Current Smoker, n (%)	14 (37)	9 (35)	5 (42)	0.73
Pack-years, yr (± SD)	43 (17)	43 (14)	53 (22)	0.26
$FEV_1\%_{pred}$ (± SD)	81 (18)	85 (13)	72 (24)	0.14
$FVC\%_{pred}$ (± SD)	82 (25)	80 (26)	86 (22)	0.62
FEV_1/FVC, % (± SD)	71 (10)	76 (5)	61 (11)	<0.0001

FEV1: Forced expiratory volume in one second, FVC: Forced vital capacity, Significance of difference ($p < 0.05$) determined using Mann Whitney t-tests for continuous variables and Fisher's exact test for categorical variables.

3.1 Inter-observer measurement reproducibility

A total of 90 OCT airway paths from all 38 subjects were acquired; 16 OCT data sets were excluded as a result of poor image quality due to motion artifacts and therefore a total of 74 data sets were included in the analysis. Table 2 shows inter-observer measurement reproducibility of all OCT-derived airway measurements for both the proximal and peripheral segments of the airway path. For all airway measurements, Pearson correlation coefficients were greater for Observer-1 (Pulmonologist) and Observer-2's (Expert) measurements, than for Observer-1 and Observer-3′s (Inexperienced Observer) measurements. Bland-Altman analysis also showed there was less bias between measurements obtained by Observer-1 and −2, than between Observer-1 and −3 for both proximal and peripheral segments, except for WT. Although the Bland-Altman bias was reduced for WT measurements obtained by Observer-1 and Observer-3 (bias = −0.09 ± 0.10mm) than for Observer-1 and Observer-2 (bias = −0.17 ± 0.10mm), there was a clear proportional error between Observer-1 and Observer-3′s measurements (Fig. 3). There was also a proportional error between Observer-2 and Observer-3′s WT measurements at both the proximal and peripheral segment of the airway path.

Table 2. Inter-Observer Reproducibility

|	Observer 1 vs. 2				Observer 1 vs. 3				Observer 2 vs. 3			
	Correlation		Bland-Altman		Correlation		Bland-Altman		Correlation		Bland-Altman	
	r	p	bias	95% CI	r	p	bias	95% CI	r	p	bias	95% CI
Proximal												
Ai (mm^2)	0.94	<0.0001	0.18 (0.16)	-0.13-0.49	0.89	<0.0001	0.18 (0.21)	-0.23-0.59	0.94	<0.0001	0.004 (0.14)	-0.27-0.27
Ao (mm^2)	0.82	<0.0001	-0.46 (0.55)	-1.54-0.61	0.79	<0.0001	0.94 (0.60)	-0.24-2.13	0.78	<0.0001	1.41 (0.59)	0.24-2.57
WA (mm^2)	0.76	<0.0001	-0.64 (0.46)	-1.56-0.27	0.66	<0.0001	0.76 (0.54)	-0.29-1.82	0.69	<0.0001	1.40 (0.58)	0.27-2.54
WT (mm)	0.76	<0.0001	-0.20 (0.12)	-0.43-0.04	0.63	<0.0001	0.21 (0.14)	-0.07-0.49	0.68	<0.0001	0.41 (0.15)	0.12-0.70
Peripheral												
Ai (mm^2)	0.38	<0.0001	0.14 (0.09)	-0.04-0.32	0.33	<0.0001	0.15 (0.10)	-0.04-0.34	0.57	<0.0001	0.006 (0.05)	-0.10-0.11
Ao (mm^2)	0.69	<0.0001	-0.35 (0.38)	-1.09-0.39	0.60	<0.0001	0.44 (0.32)	-0.18-1.07	0.59	<0.0001	0.79 (0.43)	-0.06-1.64
WA (mm^2)	0.73	<0.0001	-0.49 (0.35)	-1.17-0.19	0.61	<0.0001	0.30 (0.30)	-0.30-0.89	0.59	<0.0001	0.79 (0.44)	-0.07-1.64
WT (mm)	0.73	<0.0001	-0.17 (0.10)	-0.37-0.04	0.60	<0.0001	0.09 (0.10)	-0.10-0.28	0.58	<0.0001	0.25 (0.13)	0.0004-0.51

A. Proximal
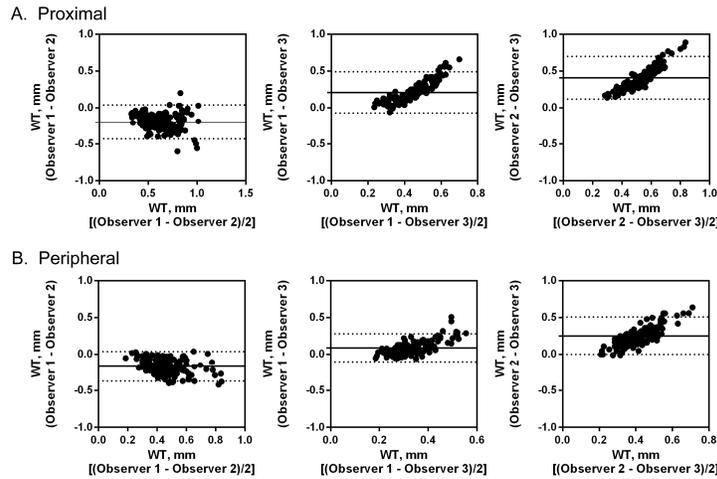
B. Peripheral

Fig. 3. Inter-Observer Reproducibility for OCT-derived WT Measurements for both Proximal and Peripheral Segments of the Airway Path.

3.2 Intra-observer measurement reproducibility

As shown in Table 3, Pearson correlation coefficients were greater for measurements obtained for segmentation round 2 and 3 than between the other segmentation rounds. However, the Bland-Altman bias was low for WT measurements between all rounds of segmentation for both proximal and peripheral measurements (Fig. 4).

Table 3. Intra-Observer Reproducibility

	Round 1 vs. 2				Round 1 vs. 3				Round 2 vs. 3			
	Correlation		Bland-Altman		Correlation		Bland-Altman		Correlation		Bland-Altman	
	r	p	bias	95% CI	r	p	bias	95% CI	r	p	bias	95% CI
Proximal												
Ai (mm^2)	0.88	<0.0001	-0.04 (0.19)	-0.41-0.33	0.89	<0.0001	-0.05 (0.19)	-0.42-0.31	0.98	<0.0001	-0.01 (0.07)	-0.16-0.14
Ao (mm^2)	0.87	<0.0001	-0.08 (0.27)	-0.61-0.46	0.86	<0.0001	-0.006 (0.28)	-0.55-0.53	0.97	<0.0001	0.07 (0.13)	-0.19-0.33
WA (mm^2)	0.74	<0.0001	-0.04 (0.15)	-0.32-0.25	0.71	<0.0001	0.05 (0.15)	-0.25-0.34	0.90	<0.0001	0.08 (0.08)	-0.08-0.25
WT (mm)	0.69	<0.0001	-0.008 (0.04)	-0.09-0.08	0.67	<0.0001	0.02 (0.04)	-0.07-0.11	0.87	<0.0001	0.03 (0.03)	-0.02-0.08
Peripheral												
Ai (mm^2)	0.35	<0.0001	-0.03 (0.10)	-0.22-0.17	0.35	<0.0001	-0.02 (0.09)	-0.20-0.17	0.49	<0.0001	0.01 (0.08)	-0.14-0.16
Ao (mm^2)	0.69	<0.0001	-0.07 (0.17)	-0.40-0.27	0.62	<0.0001	0.01 (0.18)	-0.33-0.36	0.78	<0.0001	0.08 (0.14)	-0.20-0.35
WA (mm^2)	0.83	<0.0001	-0.04 (0.10)	-0.23-0.16	0.78	<0.0001	0.03 (0.11)	-0.18-0.24	0.85	<0.0001	0.07 (0.10)	-0.12-0.26
WT (mm)	0.85	<0.0001	-0.01 (0.07)	-0.07-0.05	0.81	<0.0001	0.01 (0.04)	-0.06-0.08	0.86	<0.0001	0.02 (0.03)	-0.04-0.09

3.3 Insertion reproducibility

Of the 74 airway paths evaluated, 47 (64%) were defined to be completely matched at repeat imaging. As shown in Fig. 5, there was significantly greater insertion reproducibility in the right lung (for all lobes) compared to the left (right = 74% vs. left = 29%, p = 0.002), and there was a complete mismatch in the left upper lobe during the reinsertion imaging procedure.
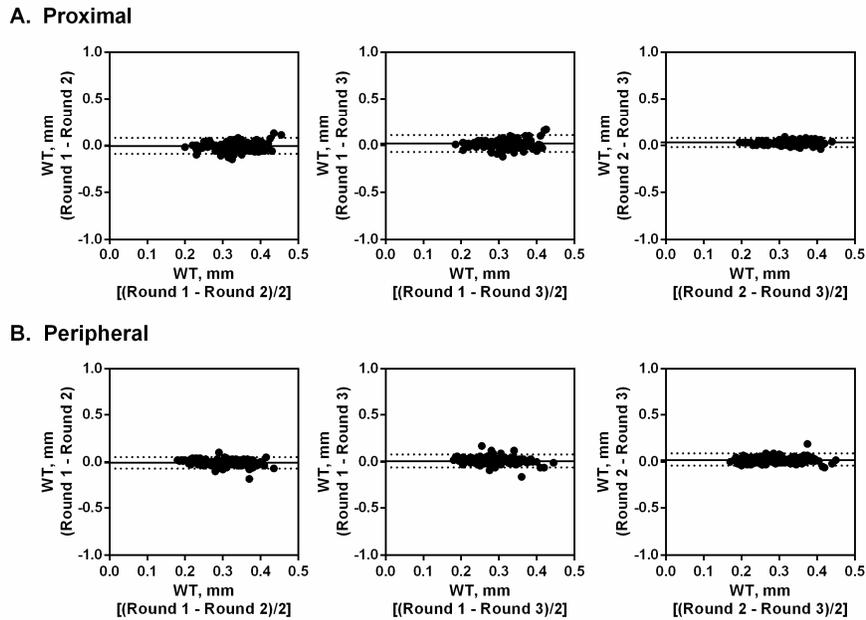

Fig. 4. Intra-Observer Reproducibility for OCT-derived WT Measurements for both Proximal and Peripheral Segments of the Airway Path.
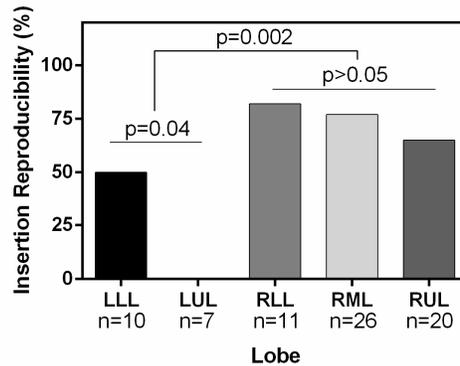

Fig. 5. Insertion-reinsertion Reproducibility. LLL = left lower lobe, LUL = left upper lobe, RLL = right lower lobe, RML = right middle lobe, RUL = right upper lobe.

3.4 OCT insertion-reinsertion measurement comparison

There were significant differences between all three observers for peripheral Ai (p = 0.04), WA (p = 0.04), WT (p = 0.02), and proximal Ao (p = 0.04), WA (p = 0.03) and WT (p =

0.02) (Table 4). However, there were no significant differences for any of the airway measurements between insertion and reinsertion (p>0.05).

Table 4. OCT Measurements for Each Observer at Insertion and Reinsertion

	Observer 1		Observer 2		Observer 3		Significance of Difference (p)	
	Insertion	Reinsertion	Insertion	Reinsertion	Insertion	Reinsertion	Observer 1-3	Insertion-reinsertion
Peripheral								
Ai (mm^2)	1.95 (0.11)	1.94 (0.09)	1.80 (0.05)	1.80 (0.05)	1.80 (0.06)	1.80 (0.06)	0.04	0.47
Ao (mm^2)	3.00 (0.40)	3.01 (0.39)	3.25 (0.53)	3.45 (0.48)	2.55 (0.19)	2.56 (0.18)	0.14	0.24
WA (mm^2)	1.04 (0.37)	1.07 (0.37)	1.44 (0.52)	1.65 (0.48)	0.76 (0.16)	0.76 (0.16)	0.04	0.28
WT (mm)	0.37 (0.12)	0.38 (0.12)	0.51 (0.16)	0.58 (0.14)	0.29 (0.05)	0.29 (0.06)	0.02	0.28
Proximal								
Ai (mm^2)	2.29 (0.41)	2.28 (0.50)	2.11 (0.38)	2.10 (0.43)	2.12 (0.38)	2.09 (0.39)	0.60	0.92
Ao (mm^2)	3.99 (0.90)	4.00 (0.95)	4.39 (0.91)	4.52 (0.89)	3.07 (0.53)	3.03 (0.51)	0.04	0.41
WA (mm^2)	1.69 (0.64)	1.72 (0.64)	2.28 (0.71)	2.42 (0.66)	0.95 (0.20)	0.94 (0.16)	0.03	0.22
WT (mm)	0.54 (0.17)	0.55 (0.17)	0.72 (0.18)	0.76 (0.17)	0.33 (0.05)	0.33 (0.04)	0.02	0.28

3.5 Insertion-reinsertion measurement reproducibility

Table 5 shows the insertion-reinsertion measurement reproducibility for all OCT-derived airway measurements for both the proximal and peripheral segments of the airway path, as well as for those airway paths that were matched or un-matched at repeat imaging. The Pearson correlation coefficients and Bland-Altman bias was similar for airway measurements that were matched and un-matched at repeat imaging. Although the 95% confidence intervals were slightly greater for un-matched WT measurements, negligible Bland-Altman bias was shown for WT measurements regardless of whether they were matched or un-matched between repeat imaging or whether the airway segment was proximal or peripherally located (Fig. 6).

Table 5. Insertion-reinsertion Measurement Reproducibility for both Proximal and Peripheral Segments of the Airway Path

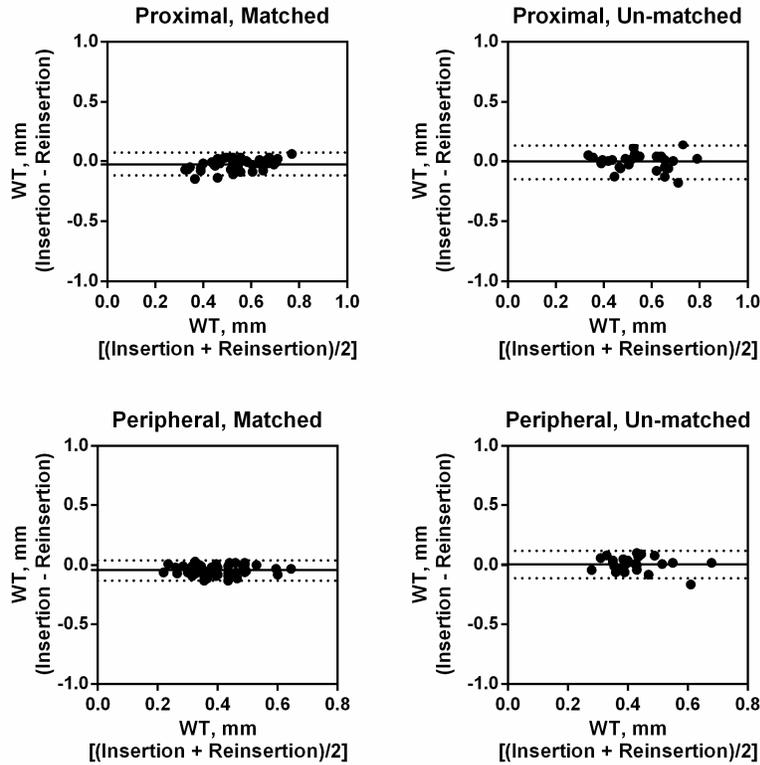|	Insertion-Reinsertion Matched					Insertion-Reinsertion Unmatched		
	Correlation			Bland-Altman			Correlation	
	r	p		r	p		r	p
Proximal			**Proximal**			**Proximal**		
Ai (mm^2)	0.81	<0.0001	Ai (mm^2)	0.81	<0.0001	Ai (mm^2)	0.81	<0.0001
Ao (mm^2)	0.91	<0.0001	Ao (mm^2)	0.91	<0.0001	Ao (mm^2)	0.91	<0.0001
WA (mm^2)	0.95	<0.0001	WA (mm^2)	0.95	<0.0001	WA (mm^2)	0.95	<0.0001
WT (mm)	0.93	<0.0001	WT (mm)	0.93	<0.0001	WT (mm)	0.93	<0.0001
Peripheral			**Peripheral**			**Peripheral**		
Ai (mm^2)	0.66	<0.0001	Ai (mm^2)	0.66	<0.0001	Ai (mm^2)	0.66	<0.0001
Ao (mm^2)	0.90	<0.0001	Ao (mm^2)	0.90	<0.0001	Ao (mm^2)	0.90	<0.0001
WA (mm^2)	0.91	<0.0001	WA (mm^2)	0.91	<0.0001	WA (mm^2)	0.91	<0.0001
WT (mm)	0.91	<0.0001	WT (mm)	0.91	<0.0001	WT (mm)	0.91	<0.0001

Fig. 6. Insertion-reinsertion Reproducibility for OCT-derived WT Measurements for both Proximal and Peripheral Matched and Un-matched Segments of the Airway Path.

Since the peripheral segments of the airway path are the likely targets for airway interventions, we provided a comparison between intra-observer and insertion-reinsertion WT measurement reproducibility calculated using the ICC and CV in Table 6. The insertion-reinsertion measurement CV was 12% (matched CV = 11%, un-matched CV = 15%) and only slightly higher than the intra-observer CV of 9%. The SDD for WT measurements was 0.08mm (matched SDD = 0.07mm, un-matched SDD = 0.09mm), and the sample sizes required for effect sizes of 10%/15%/20% (which corresponds to changes of magnitude 0.04mm, 0.06mm, 0.08mm) for all airway segments evaluated were 29/13/7.

Table 6. Comparison between Intra-Observer and Insertion-reinsertion Reproducibility for Peripheral WT Measurements

Reproducibility	Estimate
Intra-observer CV (%)	9
Intra-observer ICC [95% CI]	0.82 [0.75-0.88]
Insertion-reinsertion CV (%)	12
Matched	11
Un-matched	15
Insertion-reinsertion ICC [95% CI]	0.69 [0.55-0.79]
Matched	0.81 [0.68-0.89]
Un-matched	0.41 [0.04-0.68]
SDD (mm)	0.08
Matched	0.07
Un-matched	0.09
Sample Size Estimates for Effect Size:	
10% (0.04 mm)	29
15% (0.06 mm)	13
20% (0.08 mm)	7

4. Discussion

OCT is a bronchoscopic imaging technique that may have very important advantages over other imaging modalities for evaluating airway remodeling over time. While CT and MRI can provide volumetric images of the whole lung, only OCT can access and evaluate the small airways – the major site of airflow limitation in COPD [18]. However, in order to confidently assign changes in airway wall structure measured over time to the actual disease process and not variations in the imaging technique or observer, the reproducibility of both the OCT image acquisition technique and measurement procedure must be established.

The first step in assessing reproducibility is to evaluate inter/intra-observer measurement reproducibility. In the present study, we demonstrated inter/intra-observer measurement reproducibility was high overall, and, not surprisingly, inter-observer measurement reproducibility was greater for more experienced observers. Light signal intensity decreases as depth penetration increases in OCT images, and therefore inexperienced observers may have difficulty distinguishing the outer wall boundary. Importantly, this may explain the proportional error observed in the Bland-Altman analysis when comparing more experienced observer's measurements with those of the less experienced observer. However, the finding that there was reduced bias between the two experienced observer's measurements suggests that reproducibility improves with increased training.

Second, the reproducibility of the technique must be established to determine whether the same airway segment can be evaluated at different time-points. Because the OCT probe is a small flexible fiberoptic probe that is passed beyond the end of the bronchoscope, it may start out in a specific and known proximal airway segment, but there is potential for the probe to end up in a different, adjacent peripheral airway if it enters a different airway at the bifurcation point. Our findings indicates that the reproducibility of the probe entering the exact same peripheral airway segment was greatest in the right lung, and this is likely because of the straighter pathway for the probe to follow into the right middle and lower lobe. In contrast, the probe must negotiate sharp turns to enter either the right or left upper lobe, and this curvature may explain some of the reduced, or non-existent, upper lobe insertion reproducibility. For future studies evaluating the small airways, targeting airways in the right lobe will yield greater repeat imaging insertion reproducibility compared to the left. Furthermore, we must acknowledge that insertion-reinsertion reproducibility could have been greatly improved by a number of approaches, and further investigation into methods to improve insertion-reinsertion reproducibility are certainly warranted. For example, advancing the OCT probe through a guide sheath marked at the proximal end to determine the distance extended from the bronchoscope may improve reproducibility.

Third, serial imaging in the same subject may not result in identical images due to technical factors, such as breathing/cardiac motion artifacts, small changes in image orientation or mucus movement between image acquisitions, and this will impact measurement reproducibility. The short-term insertion-reinsertion measurement reproducibility in peripheral airways reported here was determined to be moderate for both matched (CV = 11%) and un-matched (CV = 15%) airway segments. To provide context, FEV_1 is the primary end-point that regulatory authorities regard as an acceptable measure to evaluate treatment efficacy for COPD patients in clinical trials [19]. Furthermore, FEV_1 is thought to be one of the most highly reproducible lung function parameters; the coefficients of variation for FEV_1 have been reported to be approximately 11% in patients with obstructive lung disease [20]. Therefore, OCT WT measurements have comparable reproducibility as established pulmonary function measurements that are used in clinical trials. Furthermore, the short-term insertion-reinsertion measurement reproducibility in peripheral airways was approximately the same as intra-observer measurement reproducibility (CV = 9%). This important finding suggests that much of the variability between repeat imaging, at least in this general disease type (mild COPD), can be attributed largely to the observer and not to the variability introduced by reinsertion of the OCT probe. This finding provides strong motivation for future studies to develop and validate automated segmentation

approaches to eliminate intra-observer variability and allow for even smaller differences to be detected in serial investigations. Furthermore, the finding that there was similar insertion-reinsertion measurement reproducibility for un-matched and matched airway segments may suggest that the heterogeneity of airway wall structure is relatively small within specific regions of the lung. Therefore, in studies evaluating disease longitudinally, complete matching of the OCT airway paths may not be necessary.

Finally, using the variability in the insertion-reinsertion measurements we were able to provide smallest detectable difference and sample size calculations to provide guidance for future studies. The current American Thoracic Society (ATS) and European Respiratory Society (ERS) guidelines state that spirometry measurements are standards for diagnosing COPD and evaluating treatment efficacy [21]. However, it is well-known that spirometry has limited sensitivity in COPD, and, as a result, large sample sizes are required to detect significant disease-related changes. Imaging the lung structure directly, however, may improve sensitivity and therefore reduce sample sizes and length of study times by a significant amount. For example, Dirksen *et. al.* [22] demonstrated in a cohort of subjects with alpha-1 anti-trypsin deficiency that to detect a treatment effect using FEV_1 would require 550 subjects whereas CT measurements of lung density would require only 130 subjects. While CT can assess the lung parenchyma serially, as was described by Dirksen and colleagues, other imaging approaches, such as OCT, are required to assess the small airways. We did not investigate whether OCT has sufficient sensitivity to detect small changes over time or in response to treatment in this study, however a critical first step to achieve this is to evaluate the reproducibility of the technique and measurements in order to provide estimates of the number of subjects that would be required for controlled studies. Our estimates that less than 30 subjects would be required to detect significant treatment effects bodes well for the use of OCT imaging in serial studies.

We acknowledge that this study was limited by several factors. First, the OCT catheter insertion reproducibility was assessed within the same study session instead of over two separate sessions. Therefore, the short-term variability that may be introduced by technical factors, or by short-term disease-related changes, between two imaging sessions was not assessed. However, it is important to first establish if the technique is reproducible at all, and our data suggests that it is possible to obtain reproducible measurements. Second, we only evaluated current or ex-smokers and airway changes may be more variable over time in diseases like asthma, and therefore the short-term reproducibility of OCT measurements in asthmatics should also be investigated. Third, several OCT images were discarded due to technical reasons and not included in the analysis, nor were they accounted for in our sample size calculations. We acknowledge that more effort should be made immediately following image acquisition to assess the image quality and repeat imaging if necessary. Finally, although the OCT airway wall thickness measurements used in this study have been recently validated using histology in porcine airways [15], the OCT measurements have not been validated using histology in human airways. We acknowledge that, in addition to evaluating the reproducibility of a measurement, evaluating measurement accuracy in comparison to the actual pathology is important and should be an aim of future research. Other investigations have, however, demonstrated that the airway wall features measured *in vivo* using OCT agree with histology of the corresponding airway segment in patients that had undergone lung resections [10]. This previous work provides strong support that measurements of airway wall thickness obtained using OCT reflect the actual airway pathologic changes in disease.

OCT technology is constantly evolving. While our study provides the framework for evaluating OCT insertion-reinsertion technique and WT measurement reproducibility, further research will be required as the field matures. Specifically, for our investigation, image acquisition commenced after we advanced the OCT probe until it fit within the airway of interest. However, in future studies that involve serial assessments over longer periods of time it may be more appropriate to advance the probe a known distance from a specific landmark, for example the carina, to ensure that the same airway segment is evaluated. Therefore, when new image acquisition protocols are introduced, evaluating the technique and measurement

reproducibility will be required. Furthermore, as the OCT probe technology advances and smaller diameter probes are introduced that are capable of evaluating the more distal airways, additional reproducibility studies will also be required. Clearly, there is considerable work to be done in the long-term before OCT will be used in specific clinical investigations, but our study is the first to begin to answer these questions and provide the framework for understanding and evaluating the sources of variability.

5. Summary and conclusions

In summary, we demonstrated that insertion-reinsertion reproducibility of the OCT probe was greatest in the right lung than the left (74% vs. 29%). We also demonstrated that while the overall insertion-reinsertion airway wall thickness measurement reproducibility was moderate in the peripheral airways (CV = 12%), much of the variability between repeat imaging was attributed to the observer (intra-observer CV = 9%). We also provided smallest detectable difference estimates and sample size calculations to help guide future serial studies evaluating airway remodeling. Taken together, these findings suggest that more targeted probe insertion within the right lung will improve insertion-reinsertion technique reproducibility, and the introduction of automated measurement approaches will greatly improve insertion-reinsertion measurement reproducibility thereby allowing even smaller differences in airway wall structure to be detected over time. In conclusion, this study demonstrates that OCT has the potential to be used in future longitudinal studies to provide a better understanding of disease pathogenesis and response to treatment.