Lung-airway Data Interrogation via Cluster Analysis

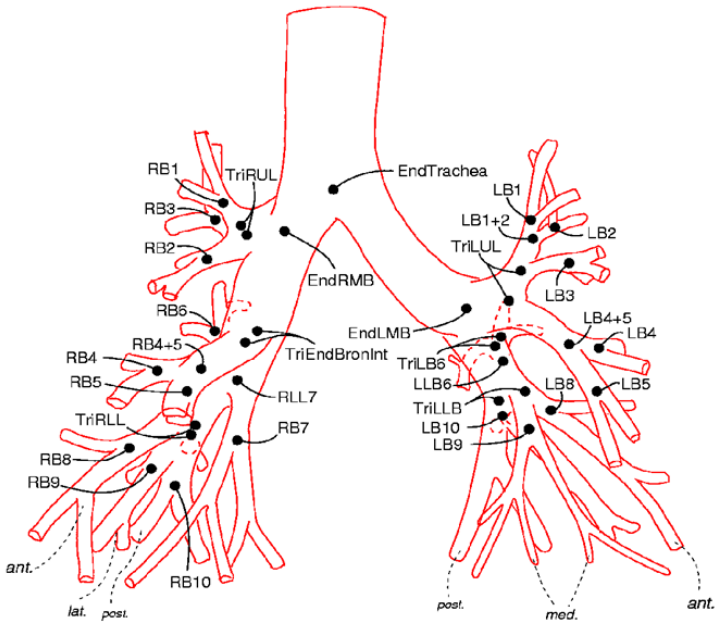
Kung-Sik Chan[1], Ching-Long Lin[1], Kun Chen[2] and Eric A. Hoffman[1]

The University of Iowa[1] and Kansas State University[2]

Oct 5, 2011

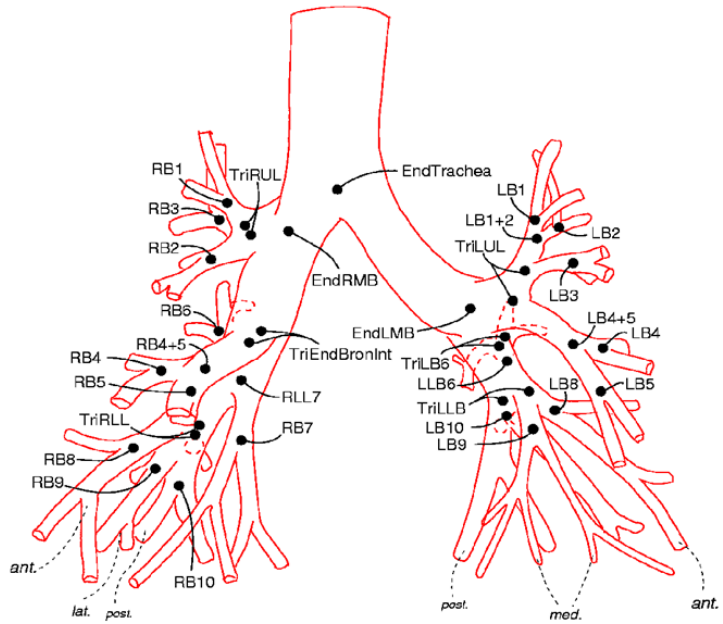
Introduction

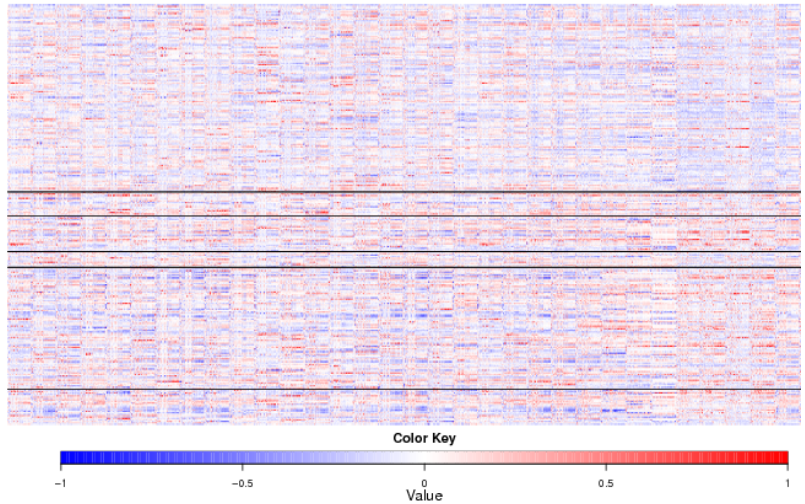
- Multi detector-row CT (MDCT)-based imaging and automated image analysis yield rather detailed *in vivo* measurements of individual human lung airways.



Introduction

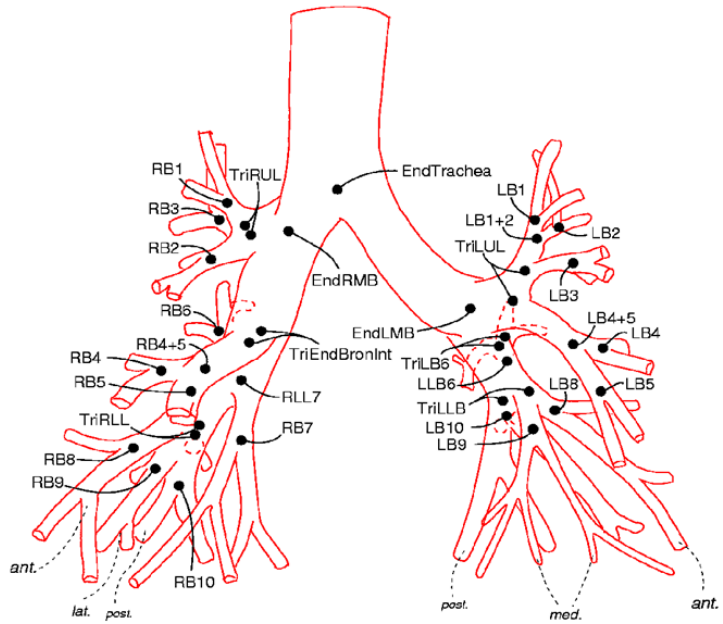
- Multi detector-row CT (MDCT)-based imaging and automated image analysis yield rather detailed *in vivo* measurements of individual human lung airways.
- Lung airway data comprise high-dimensional image variables, e.g. segment-by-segment average wall thickness, average inner area, branch angles, etc.

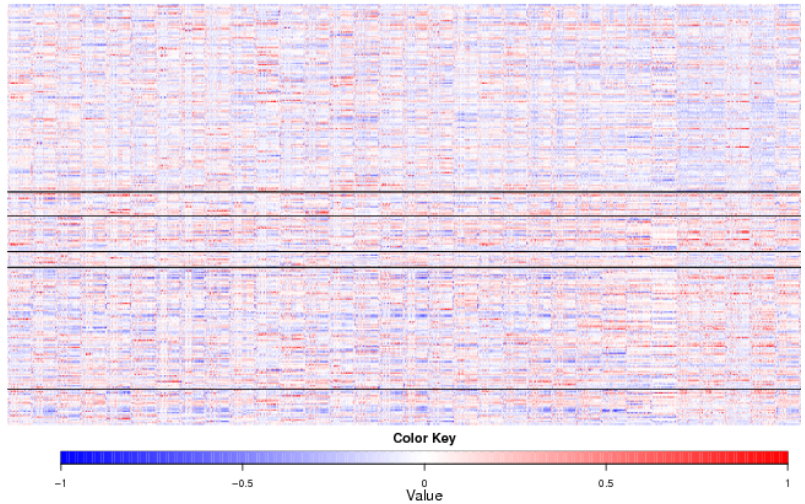




Introduction

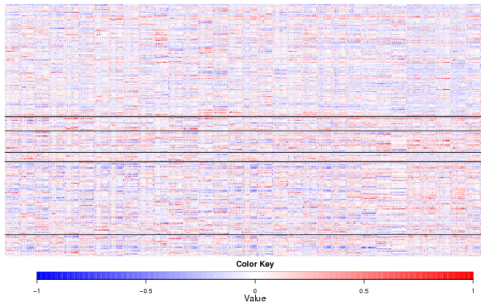
- Multi detector-row CT (MDCT)-based imaging and automated image analysis yield rather detailed *in vivo* measurements of individual human lung airways.
- Lung airway data comprise high-dimensional image variables, e.g. segment-by-segment average wall thickness, average inner area, branch angles, etc.
- Issue: explore the presence of systematic variations in the lung airway structure between normal non-smokers (n=132) and asthmatics from SARP (n=164). 1567 (derived) airway measurements per subject.





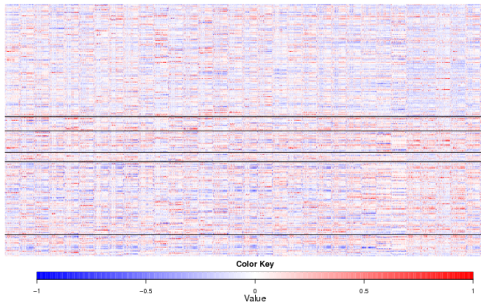
Method

$$\begin{pmatrix} Y_{1,1} & Y_{1,2} & \cdots & Y_{1,1567} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{132,1} & Y_{132,2} & \cdots & Y_{132,1567} \\ Y_{133,1} & Y_{133,2} & \cdots & Y_{133,1567} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{296,1} & Y_{296,2} & \cdots & Y_{296,1567} \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{164}{132 \times 296}} \\ \vdots \\ \sqrt{\frac{164}{132 \times 296}} \\ -\sqrt{\frac{132}{164 \times 296}} \\ \vdots \\ -\sqrt{\frac{132}{164 \times 296}} \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_{1567} \end{pmatrix} + E$$



Method

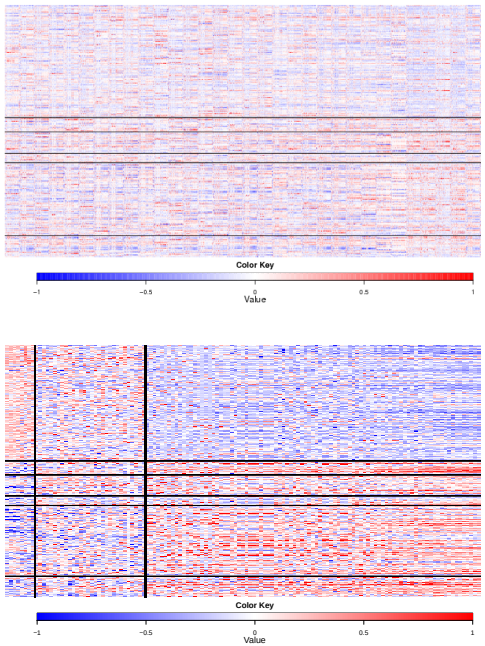
$$\begin{pmatrix} Y_{1,1} & Y_{1,2} & \cdots & Y_{1,1567} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{132,1} & Y_{132,2} & \cdots & Y_{132,1567} \\ Y_{133,1} & Y_{133,2} & \cdots & Y_{133,1567} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{296,1} & Y_{296,2} & \cdots & Y_{296,1567} \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{164}{132 \times 296}} \\ \vdots \\ \sqrt{\frac{164}{132 \times 296}} \\ -\sqrt{\frac{132}{164 \times 296}} \\ \vdots \\ -\sqrt{\frac{132}{164 \times 296}} \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_{1567} \end{pmatrix} + E$$



- Non-zero v's correspond to important variables distinguishing normal non-smokers from asthmatics.

Method

$$\begin{pmatrix} Y_{1,1} & Y_{1,2} & \cdots & Y_{1,1567} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{132,1} & Y_{132,2} & \cdots & Y_{132,1567} \\ Y_{133,1} & Y_{133,2} & \cdots & Y_{133,1567} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{296,1} & Y_{296,2} & \cdots & Y_{296,1567} \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{164}{132 \times 296}} \\ \vdots \\ \sqrt{\frac{164}{132 \times 296}} \\ -\sqrt{\frac{132}{164 \times 296}} \\ \vdots \\ -\sqrt{\frac{132}{164 \times 296}} \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_{1567} \end{pmatrix} + E$$



- Non-zero v's correspond to important variables distinguishing normal non-smokers from asthmatics.
- Employ a new statistical method (RRR-SVD) that provides good fit with a sparse estimator of the v's.

Results

Out of the 100 leading variables selected by RRR-SVD

- 59 variables concern wall thickness, mostly of branches of generations 3-7.

Results

Out of the 100 leading variables selected by RRR-SVD

- 59 variables concern wall thickness, mostly of branches of generations 3-7.
- 17 variables are related to the **branch angles** between adjacent segments.

Results

Out of the 100 leading variables selected by RRR-SVD

- 59 variables concern wall thickness, mostly of branches of generations 3-7.
- 17 variables are related to the **branch angles** between adjacent segments.
- 24 variables concern the variability of measurements across generations.

Results

Out of the 100 leading variables selected by RRR-SVD

- 59 variables concern wall thickness, mostly of branches of generations 3-7.
- 17 variables are related to the **branch angles** between adjacent segments.
- 24 variables concern the variability of measurements across generations.

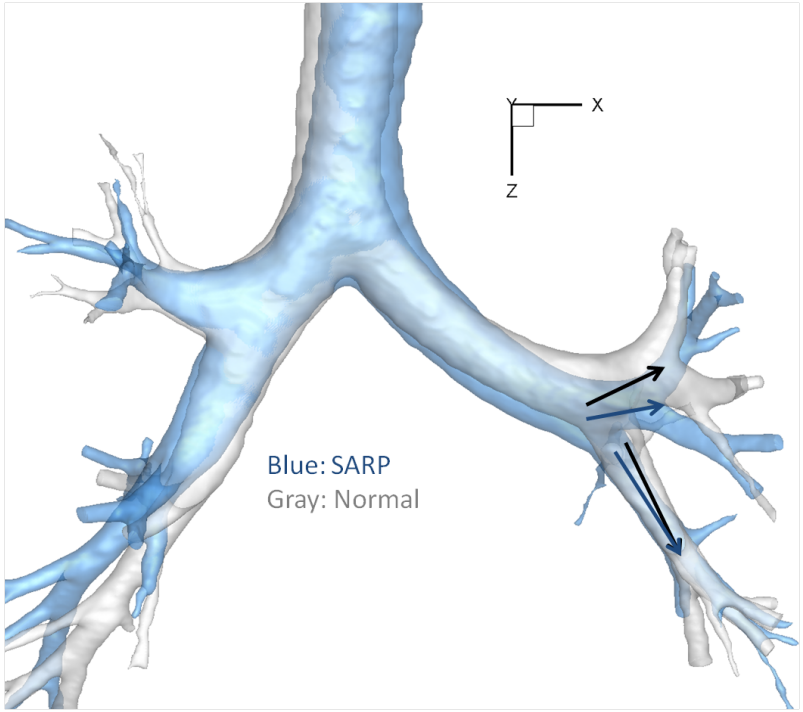
Wall thickness is an indicator for airway remodeling and rigidity, the branch angle difference may be associated with lung inflation and tissue structure, and variability across generations may imply, e.g. airway constriction.

Results

Out of the 100 leading variables selected by RRR-SVD

- 59 variables concern wall thickness, mostly of branches of generations 3-7.
- 17 variables are related to the **branch angles** between adjacent segments.
- 24 variables concern the variability of measurements across generations.

Wall thickness is an indicator for airway remodeling and rigidity, the branch angle difference may be associated with lung inflation and tissue structure, and variability across generations may imply, e.g. airway constriction.

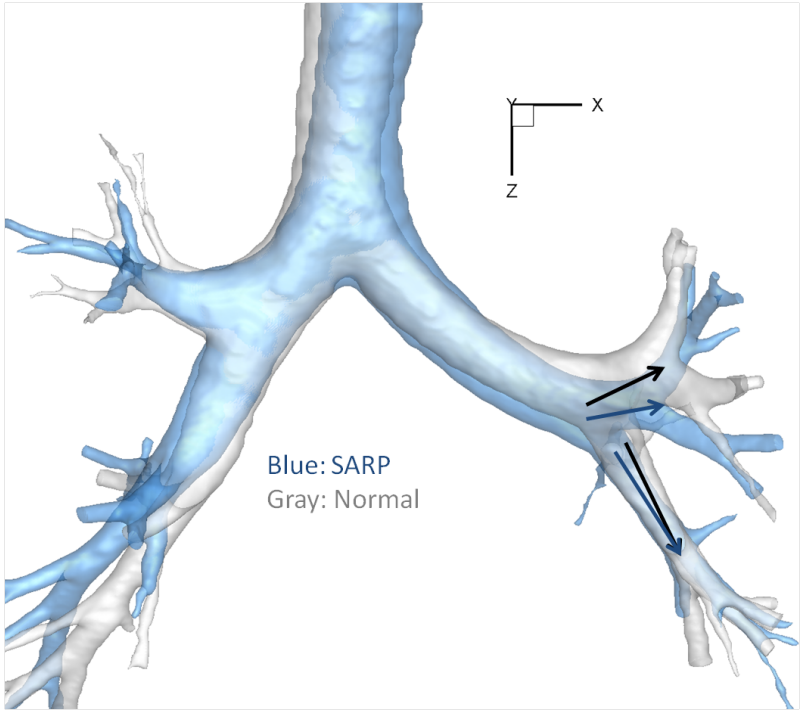


Results

Out of the 100 leading variables selected by RRR-SVD

- 59 variables concern wall thickness, mostly of branches of generations 3-7.
- 17 variables are related to the **branch angles** between adjacent segments.
- 24 variables concern the variability of measurements across generations.

Wall thickness is an indicator for airway remodeling and rigidity, the branch angle difference may be associated with lung inflation and tissue structure, and variability across generations may imply, e.g. airway constriction.



	Out-of-sample		
	MCR	FPR	FNR
With all variables	12.2%	10.6%	13.4%
Without "angle" variables	15.5%	17.4%	14.0%
"Angle" variables only	16.2%	18.2%	14.6%

MCR: misclassification rate
FPR: rate of misclassifying normal subjects as asthmatics
FNR: rate of misclassifying asthmatics as normal
These rates are based on twenty-fold cross validation

Thank You!

Thank You!

We gratefully acknowledge the support by NIH Grants R01-HL-064368 and R01-HL-094315.