Gender difference in pulmonary functions and lung parenchyma measurements: The use of propensity scores

Jorge León Cruz
Paula Rivera Torres

University of Puerto Rico, Cayey, PR
University of Puerto Rico, Humacao, PR

July 27, 2011

Talk Outline

Outline

- Background
- Goal
- Method
- Data Analysis / Results
- Conclusion / Future Work / References

Talk Outline

Outline

- Background
- Goal
- Method
- Data Analysis / Results
- Conclusion / Future Work / References

Talk Outline

Outline

- Background
- Goal
- Method
- Data Analysis / Results
- Conclusion / Future Work / References

Talk Outline

Outline

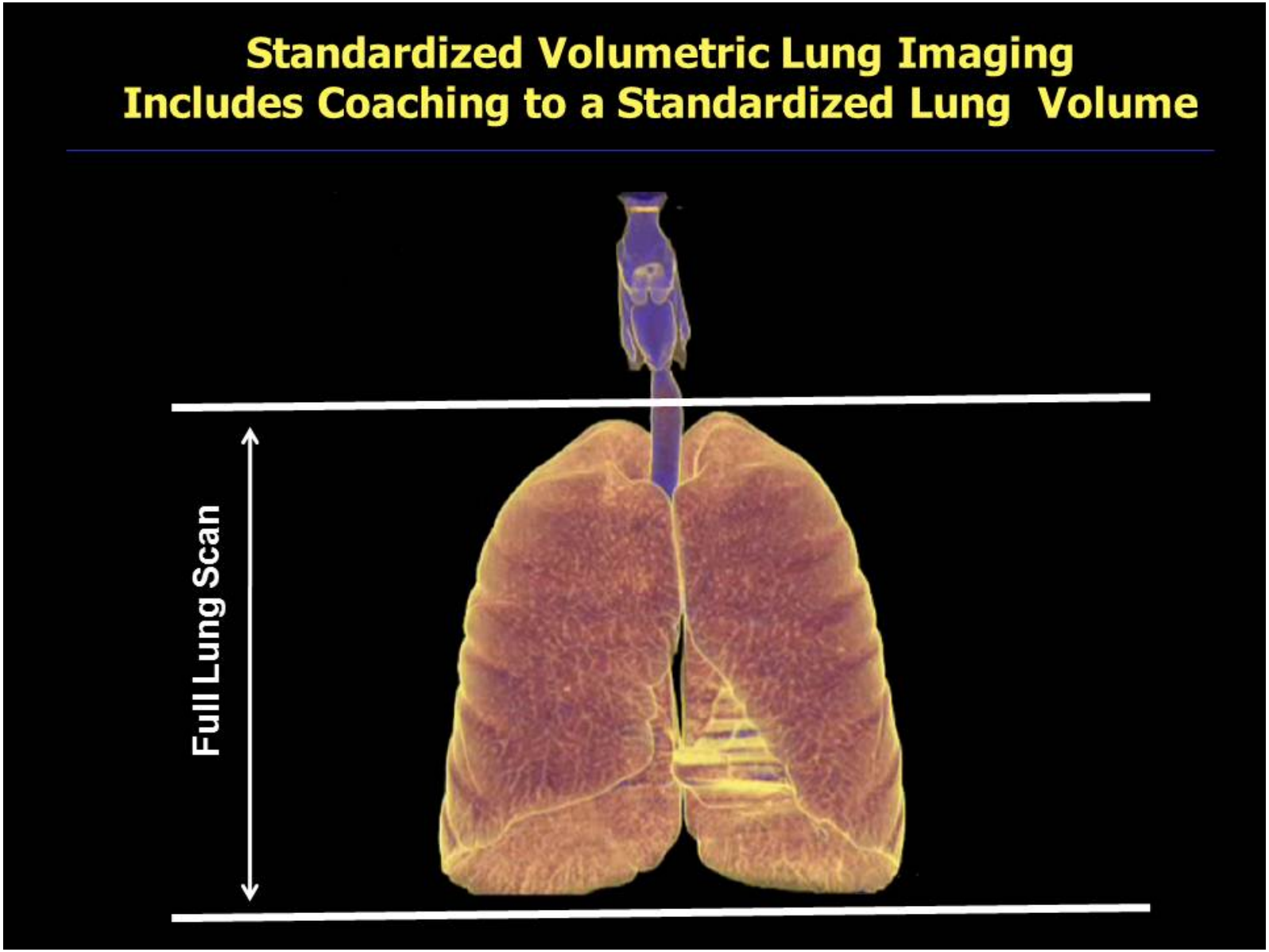
- Background
- Goal
- Method
- Data Analysis / Results
- Conclusion / Future Work / References

Talk Outline

Outline

- Background
- Goal
- Method
- Data Analysis / Results
- Conclusion / Future Work / References

Background


Standardized Volumetric Lung Imaging
Includes Coaching to a Standardized Lung Volume
Full Lung Scan

Pulmonary Function Testing (Defns.)

- **Forced Vital Capacity (FVC)**. Maximum volume of air that can be forcibly expired after full inspiration
- **Forced Expired Volume in 1 Second (FEV1)**. Measured after maximum inspiration, the volume of air that can be expelled in 1 s
- **Residual Volume (RV)**. The volume of air remaining in the lungs at the end of a maximal exhalation
- **Total Lung Capacity (TLC)**. Volume of air in the lungs after maximal inspiration
- **Functional Residual Capacity (FRC)**. The amount of gas remaining in the lung after a normal tidal volume expiration.
- These measures represent the integrated state of the lung function.
- To study the region state of the lung function we turn to computed tomography (CT).

Pulmonary Function Testing (Defns.)

- **Forced Vital Capacity (FVC)**. Maximum volume of air that can be forcibly expired after full inspiration
- **Forced Expired Volume in 1 Second (FEV1)**. Measured after maximum inspiration, the volume of air that can be expelled in 1 s
- **Residual Volume (RV)**. The volume of air remaining in the lungs at the end of a maximal exhalation
- **Total Lung Capacity (TLC)**. Volume of air in the lungs after maximal inspiration
- **Functional Residual Capacity (FRC)**. The amount of gas remaining in the lung after a normal tidal volume expiration.
- These measures represent the integrated state of the lung function.
- To study the region state of the lung function we turn to computed tomography (CT).

Pulmonary Function Testing (Defns.)

- **Forced Vital Capacity (FVC)**. Maximum volume of air that can be forcibly expired after full inspiration
- **Forced Expired Volume in 1 Second (FEV1)**. Measured after maximum inspiration, the volume of air that can be expelled in 1 s
- **Residual Volume (RV)**. The volume of air remaining in the lungs at the end of a maximal exhalation
- **Total Lung Capacity (TLC)**. Volume of air in the lungs after maximal inspiration
- **Functional Residual Capacity (FRC)**. The amount of gas remaining in the lung after a normal tidal volume expiration.
- These measures represent the integrated state of the lung function.
- To study the region state of the lung function we turn to computed tomography (CT).

Pulmonary Function Testing (Defns.)

- **Forced Vital Capacity (FVC)**. Maximum volume of air that can be forcibly expired after full inspiration
- **Forced Expired Volume in 1 Second (FEV1)**. Measured after maximum inspiration, the volume of air that can be expelled in 1 s
- **Residual Volume (RV)**. The volume of air remaining in the lungs at the end of a maximal exhalation
- **Total Lung Capacity (TLC)**. Volume of air in the lungs after maximal inspiration
- **Functional Residual Capacity (FRC)**. The amount of gas remaining in the lung after a normal tidal volume expiration.
- These measures represent the integrated state of the lung function.
- To study the region state of the lung function we turn to computed tomography (CT).

Pulmonary Function Testing (Defns.)

- **Forced Vital Capacity (FVC)**. Maximum volume of air that can be forcibly expired after full inspiration
- **Forced Expired Volume in 1 Second (FEV1)**. Measured after maximum inspiration, the volume of air that can be expelled in 1 s
- **Residual Volume (RV)**. The volume of air remaining in the lungs at the end of a maximal exhalation
- **Total Lung Capacity (TLC)**. Volume of air in the lungs after maximal inspiration
- **Functional Residual Capacity (FRC)**. The amount of gas remaining in the lung after a normal tidal volume expiration.
- These measures represent the integrated state of the lung function.
- To study the region state of the lung function we turn to computed tomography (CT).

Pulmonary Function Testing (Defns.)

- **Forced Vital Capacity (FVC)**. Maximum volume of air that can be forcibly expired after full inspiration
- **Forced Expired Volume in 1 Second (FEV1)**. Measured after maximum inspiration, the volume of air that can be expelled in 1 s
- **Residual Volume (RV)**. The volume of air remaining in the lungs at the end of a maximal exhalation
- **Total Lung Capacity (TLC)**. Volume of air in the lungs after maximal inspiration
- **Functional Residual Capacity (FRC)**. The amount of gas remaining in the lung after a normal tidal volume expiration.
- These measures represent the integrated state of the lung function.
- To study the region state of the lung function we turn to computed tomography (CT).

Pulmonary Function Testing (Defns.)

- **Forced Vital Capacity (FVC)**. Maximum volume of air that can be forcibly expired after full inspiration
- **Forced Expired Volume in 1 Second (FEV1)**. Measured after maximum inspiration, the volume of air that can be expelled in 1 s
- **Residual Volume (RV)**. The volume of air remaining in the lungs at the end of a maximal exhalation
- **Total Lung Capacity (TLC)**. Volume of air in the lungs after maximal inspiration
- **Functional Residual Capacity (FRC)**. The amount of gas remaining in the lung after a normal tidal volume expiration.
- These measures represent the integrated state of the lung function.
- To study the region state of the lung function we turn to computed tomography (CT).

Lung Parenchyma measurement

- Obtained from computed tomography (CT)
- Lung density histogram is calculated from relative attenuation differences in the medium.
- Hounsfield Units are used to describe the relative attenuation where air is defined as $-1000HU$ and water as $0HU$.
- Low attenuation areas ($< -856HU$) indicate the structure of the lung at the terminal level. (measured during FRC)
- $-910HU$ and $-950HU$ are associated with ephysema progression (measured during TLC)
- The high resolution and definition of the CT scans allow to create detailed models of the entire lung structure.

Lung Parenchyma measurement

- Obtained from computed tomography (CT)
- Lung density histogram is calculated from relative attenuation differences in the medium.
- Hounsfield Units are used to describe the relative attenuation where air is defined as $-1000HU$ and water as $0HU$.
- Low attenuation areas ($< -856HU$) indicate the structure of the lung at the terminal level. (measured during FRC)
- $-910HU$ and $-950HU$ are associated with ephysema progression (measured during TLC)
- The high resolution and definition of the CT scans allow to create detailed models of the entire lung structure.

Lung Parenchyma measurement

- Obtained from computed tomography (CT)
- Lung density histogram is calculated from relative attenuation differences in the medium.
- Hounsfield Units are used to describe the relative attenuation where air is defined as $-1000HU$ and water as $0HU$.
- Low attenuation areas ($< -856HU$) indicate the structure of the lung at the terminal level. (measured during FRC)
- $-910HU$ and $-950HU$ are associated with ephysema progression (measured during TLC)
- The high resolution and definition of the CT scans allow to create detailed models of the entire lung structure.

Lung Parenchyma measurement

- Obtained from computed tomography (CT)
- Lung density histogram is calculated from relative attenuation differences in the medium.
- Hounsfield Units are used to describe the relative attenuation where air is defined as $-1000HU$ and water as $0HU$.
- Low attenuation areas ($< -856HU$) indicate the structure of the lung at the terminal level. (measured during FRC)
- $-910HU$ and $-950HU$ are associated with ephysema progression (measured during TLC)
- The high resolution and definition of the CT scans allow to create detailed models of the entire lung structure.

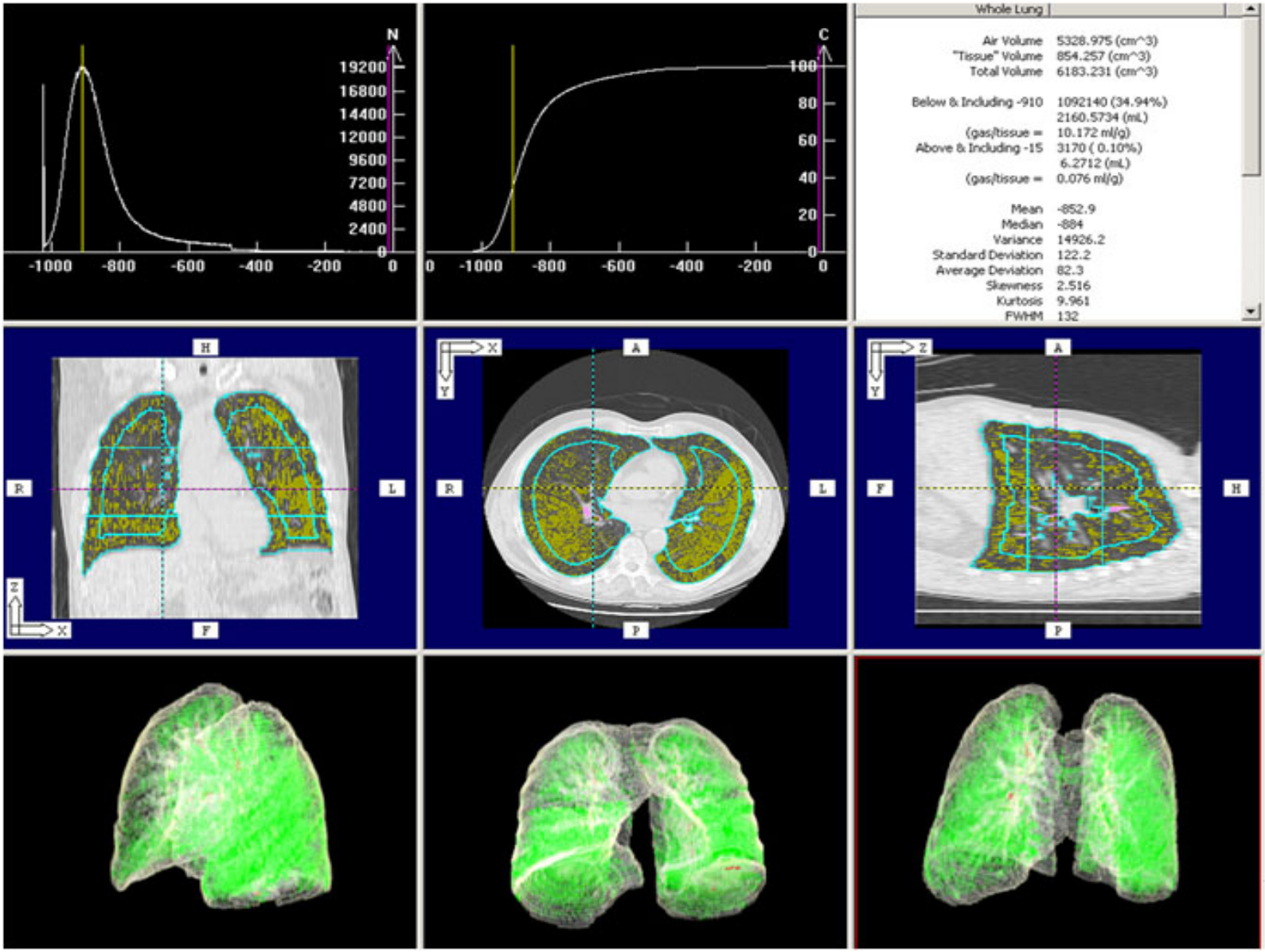
Lung Parenchyma measurement

- Obtained from computed tomography (CT)
- Lung density histogram is calculated from relative attenuation differences in the medium.
- Hounsfield Units are used to describe the relative attenuation where air is defined as $-1000HU$ and water as $0HU$.
- Low attenuation areas ($< -856HU$) indicate the structure of the lung at the terminal level. (measured during FRC)
- $-910HU$ and $-950HU$ are associated with ephysema progression (measured during TLC)
- The high resolution and definition of the CT scans allow to create detailed models of the entire lung structure.

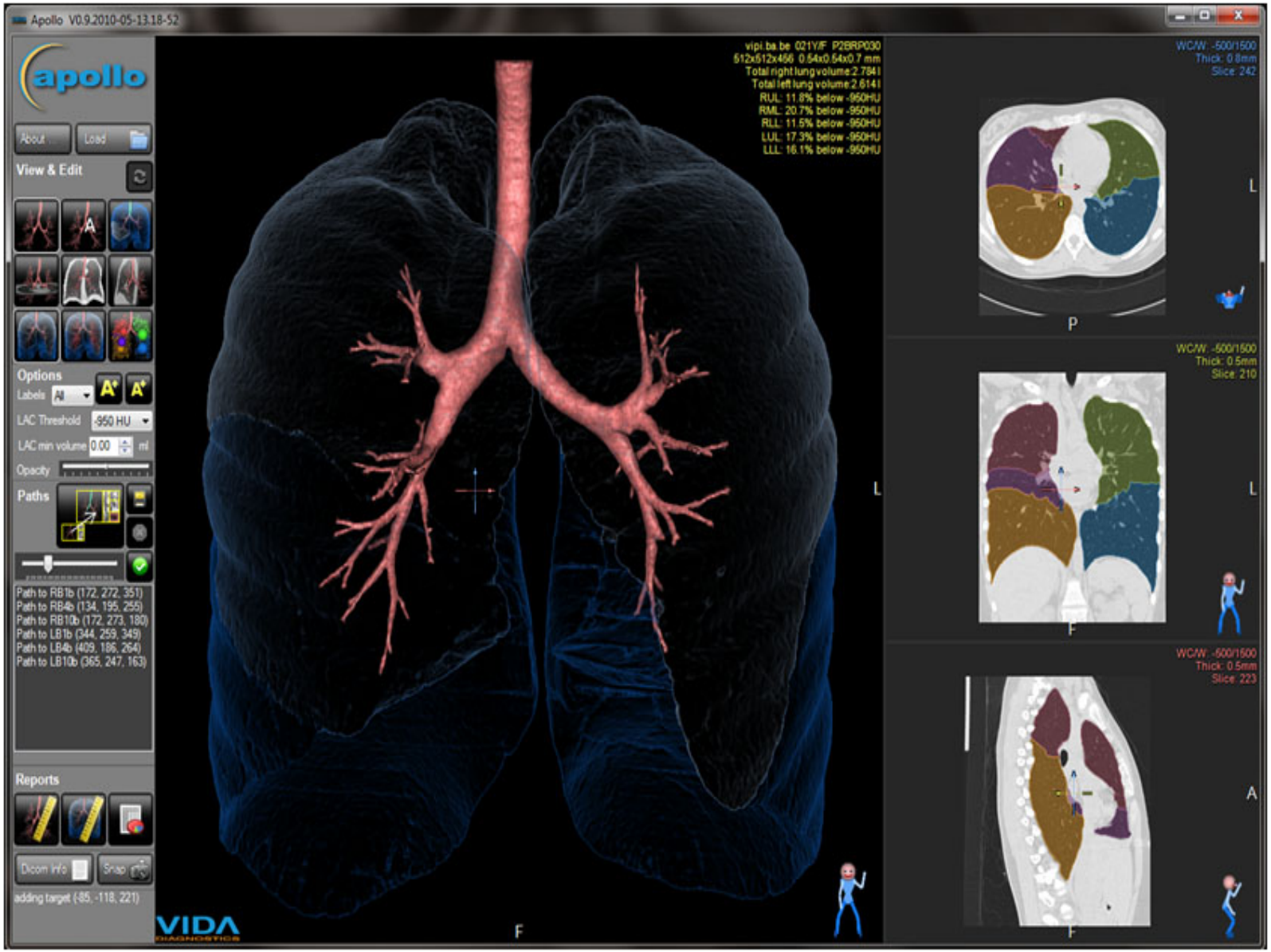
Lung Parenchyma measurement

- Obtained from computed tomography (CT)
- Lung density histogram is calculated from relative attenuation differences in the medium.
- Hounsfield Units are used to describe the relative attenuation where air is defined as $-1000HU$ and water as $0HU$.
- Low attenuation areas ($< -856HU$) indicate the structure of the lung at the terminal level. (measured during FRC)
- $-910HU$ and $-950HU$ are associated with ephysema progression (measured during TLC)
- The high resolution and definition of the CT scans allow to create detailed models of the entire lung structure.

Background


N
19200
16800
14400
12000
9600
7200
4800
2400
0
-1000
-800
-600
-400
-200
0
C
100
80
60
40
20
0
Whole Lung
Air Volume 5328.975 (cm^3)
"Tissue" Volume 854.257 (cm^3)
Total Volume 6183.231 (cm^3)
Below & Including -910 1092140 (34.94%)
2160.5734 (mL)
(gas/tissue = 10.172 ml/g)
Above & Including -15 3170 (0.10%)
6.2712 (mL)
(gas/tissue = 0.076 ml/g)
Mean -852.9
Median -884
Variance 14926.2
Standard Deviation 122.2
Average Deviation 82.3
Skewness 2.516
Kurtosis 9.961
FWHM 132
H
R
L
F
A
P
Z
X
Y

Background


Apollo V0.9.2010-05-13.18-52
apollo
About
Load
View & Edit
Options
Labels
LAC Threshold
-950 HU
LAC min volume
0.00
Opacity
Paths
Path to RB1b (172, 272, 351)
Path to RB4b (134, 195, 255)
Path to RB10b (172, 273, 180)
Path to LB1b (344, 259, 349)
Path to LB4b (409, 186, 264)
Path to LB10b (365, 247, 163)
Reports
Dicom Info
Snap
adding target (-85, -118, 221)
VIDA
vipi.ba.be 021Y/F P2BRP030
512x512x456 0.64x0.64x0.7 mm
Total right lung volume:2.784 l
Total left lung volume:2.614 l
RUL: 11.8% below -950HU
RML: 20.7% below -950HU
RLL: 11.5% below -950HU
LUL: 17.3% below -950HU
LLL: 16.1% below -950HU
WC/W: -500/1500
Thick: 0.8mm
Slice: 242
WC/W: -500/1500
Thick: 0.5mm
Slice: 210
WC/W: -500/1500
Thick: 0.5mm
Slice: 223

Background

- As expected, male lungs are on average larger than female lungs.
- Also, each respective gender pulmonary function and lung parenchyma measures are different.
- For example,

Baseline Normal Characteristic by gender			
Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
Pulmonary functions measure:			
rv/tlc	0.29	0.34	0.004
Lung parenchyma measure:			
total volume	6914.10	5218.56	2.2×10^{-16}

- Are there fundamental biological differences in male and female lungs?
- Or is this difference due to the size difference between genders?

Background

- As expected, male lungs are on average larger than female lungs.
- Also, each respective gender pulmonary function and lung parenchyma measures are different.
- For example,

Baseline Normal Characteristic by gender

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
Pulmonary functions measure:			
rv/tlc	0.29	0.34	0.004
Lung parenchyma measure:			
total volume	6914.10	5218.56	2.2×10^{-16}

- Are there fundamental biological differences in male and female lungs?
- Or is this difference due to the size difference between genders?

Background

- As expected, male lungs are on average larger than female lungs.
- Also, each respective gender pulmonary function and lung parenchyma measures are different.
- For example,

Baseline Normal Characteristic by gender			
Variable	Male (N = 51)	Female (N = 74)	P-value
	Mean	Mean	
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
Pulmonary functions measure:			
rv/tlc	0.29	0.34	0.004
Lung parenchyma measure:			
total volume	6914.10	5218.56	2.2×10^{-16}

- Are there fundamental biological differences in male and female lungs?
- Or is this difference due to the size difference between genders?

Background

- As expected, male lungs are on average larger than female lungs.
- Also, each respective gender pulmonary function and lung parenchyma measures are different.
- For example,

Baseline Normal Characteristic by gender			
Variable	Male (N = 51)	Female (N = 74)	P-value
	Mean	Mean	
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
Pulmonary functions measure:			
rv/tlc	0.29	0.34	0.004
Lung parenchyma measure:			
total volume	6914.10	5218.56	2.2×10^{-16}

- Are there fundamental biological differences in male and female lungs?
- Or is this difference due to the size difference between genders?

Background

- As expected, male lungs are on average larger than female lungs.
- Also, each respective gender pulmonary function and lung parenchyma measures are different.
- For example,

Baseline Normal Characteristic by gender			
Variable	Male (N = 51)	Female (N = 74)	P-value
	Mean	Mean	
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
Pulmonary functions measure:			
rv/tlc	0.29	0.34	0.004
Lung parenchyma measure:			
total volume	6914.10	5218.56	2.2×10^{-16}

- Are there fundamental biological differences in male and female lungs?
- Or is this difference due to the size difference between genders?

Goal

Goal

Our research project explores these differences to investigate whether size confounds the gender differences or whether other biological variables may be at stake. The project uses propensity scores for this exploration by matching males and females based on physical characteristics to remove potential bias due to size, and assess gender differences in pulmonary function and lung parenchyma on the matched subset.

Propensity Score

Definition

Propensity score is the conditional probability of being treated given covariates.

Uses

- balances covariates in the two groups (treated and non-treated),
- reduces bias estimates of treatment effects, and
- increases precision in studies.

Propensity Score

Definition

Propensity score is the conditional probability of being treated given covariates.

Uses

- balances covariates in the two groups (treated and non-treated),
- reduces bias estimates of treatment effects, and
- increases precision in studies.

Propensity Score

Definition

Propensity score is the conditional probability of being treated given covariates.

Uses

- balances covariates in the two groups (treated and non-treated),
- reduces bias estimates of treatment effects, and
- increases precision in studies.

Statistical Definition

Definition

The estimated *propensity score* , for subject i, $(i = 1, ..., N)$ is the conditional probability of assignment to a particular treatment $(Z_i = 1)$ versus control $(Z_i = 0)$ given a vector of observed covariates X_i (Rosenbaum and Rubin, 1998):

$$e(x_i) = Pr(Z_i = 1|X_i = x_i)$$

where it is assumed that, given the $X's$, the Z_i are independent:

$$Pr(Z_1 = z_1, ..., Z_N = z_N|X_1 = x_1, ..., X_N = x_N) =$$

$$\prod_{i=1}^{N} e(x_i)^{z_i} [1 - e(x_i)]^{1-z_i}$$

Method for Calculating Propensity Scores

Logistic Regression

- model used to predict the probability that an event occurs
- $P_i = ln\left(\frac{e(x_i)}{1-e(x_i)}\right) = ln\left(\frac{Pr(z_i=1|x_i)}{1-Pr(z_i=1|x_i)}\right) = \alpha + \beta^T X_i$
  where:
- X_i is a vector of the observed covariates
- β is a vector of the estimated regression coefficients computed using maximum likelihood estimation.

Method for Calculating Propensity Scores

Logistic Regression

- model used to predict the probability that an event occurs
- $P_i = ln\left(\frac{e(x_i)}{1-e(x_i)}\right) = ln\left(\frac{Pr(z_i=1|x_i)}{1-Pr(z_i=1|x_i)}\right) = \alpha + \beta^T X_i$
  where:
- X_i is a vector of the observed covariates
- β is a vector of the estimated regression coefficients computed using maximum likelihood estimation.

Method for Calculating Propensity Scores

Logistic Regression

- model used to predict the probability that an event occurs
- $P_i = ln\left(\frac{e(x_i)}{1-e(x_i)}\right) = ln\left(\frac{Pr(z_i=1|x_i)}{1-Pr(z_i=1|x_i)}\right) = \alpha + \beta^T X_i$
  where:
- X_i is a vector of the observed covariates
- β is a vector of the estimated regression coefficients computed using maximum likelihood estimation.

Method for Calculating Propensity Scores

Logistic Regression

- model used to predict the probability that an event occurs
- $P_i = ln\left(\frac{e(x_i)}{1-e(x_i)}\right) = ln\left(\frac{Pr(z_i=1|x_i)}{1-Pr(z_i=1|x_i)}\right) = \alpha + \beta^T X_i$
  where:
- X_i is a vector of the observed covariates
- β is a vector of the estimated regression coefficients computed using maximum likelihood estimation.

Matching

- Matching is a common technique used to select control subject who are matched with the treated subjects on background covariates that the investigator believes need to be controlled.
- Although the idea of finding matches seems straightforward, it is often difficult to find subjects who are similar on all important covariates.
- Some methods used to match subjects are: *Matching by the nearest available logit of the propensity scores*, *Mahalanobis metric matching*, and a combination of both.

Matching

- Matching is a common technique used to select control subject who are matched with the treated subjects on background covariates that the investigator believes need to be controlled.
- Although the idea of finding matches seems straightforward, it is often difficult to find subjects who are similar on all important covariates.
- Some methods used to match subjects are: *Matching by the nearest available logit of the propensity scores*, *Mahalanobis metric matching*, and a combination of both.

Matching

- Matching is a common technique used to select control subject who are matched with the treated subjects on background covariates that the investigator believes need to be controlled.
- Although the idea of finding matches seems straightforward, it is often difficult to find subjects who are similar on all important covariates.
- Some methods used to match subjects are: *Matching by the nearest available logit of the propensity scores*, *Mahalanobis metric matching*, and a combination of both.

Matching by the nearest Logit of the Propensity Scores within tolerance level (one-to-one)

$m(P_i) = min_j \{|P_i - P_j| < \epsilon\}$

- Step 1: Calculate the estimated propensity scores for treated and non-treated samples.
- Step 2: Matches for treated subject i are selected, if possible, only if $|P_i - P_j| < \epsilon$, where ϵ is a pre-specified tolerance.
- Step 3:The control subject with the value of P_j that is closest to P_i is selected as the match, and both are removed from the pool.
- Step 4: Repeat Step 2 and 3 until matching is no longer possible.
- Step 5: Discard the unmatched subjects.

Matching by the nearest Logit of the Propensity Scores within tolerance level (one-to-one)

$m(P_i) = min_j \{|P_i - P_j| < \epsilon\}$

- Step 1: Calculate the estimated propensity scores for treated and non-treated samples.
- Step 2: Matches for treated subject i are selected, if possible, only if $|P_i - P_j| < \epsilon$, where ϵ is a pre-specified tolerance.
- Step 3:The control subject with the value of P_j that is closest to P_i is selected as the match, and both are removed from the pool.
- Step 4: Repeat Step 2 and 3 until matching is no longer possible.
- Step 5: Discard the unmatched subjects.

Matching by the nearest Logit of the Propensity Scores within tolerance level (one-to-one)

$m(P_i) = min_j \{|P_i - P_j| < \epsilon\}$

- Step 1: Calculate the estimated propensity scores for treated and non-treated samples.
- Step 2: Matches for treated subject i are selected, if possible, only if $|P_i - P_j| < \epsilon$, where ϵ is a pre-specified tolerance.
- Step 3:The control subject with the value of P_j that is closest to P_i is selected as the match, and both are removed from the pool.
- Step 4: Repeat Step 2 and 3 until matching is no longer possible.
- Step 5: Discard the unmatched subjects.

Matching by the nearest Logit of the Propensity Scores within tolerance level (one-to-one)

$m(P_i) = min_j \{|P_i - P_j| < \epsilon\}$

- Step 1: Calculate the estimated propensity scores for treated and non-treated samples.
- Step 2: Matches for treated subject i are selected, if possible, only if $|P_i - P_j| < \epsilon$, where ϵ is a pre-specified tolerance.
- Step 3:The control subject with the value of P_j that is closest to P_i is selected as the match, and both are removed from the pool.
- Step 4: Repeat Step 2 and 3 until matching is no longer possible.
- Step 5: Discard the unmatched subjects.

Matching by the nearest Logit of the Propensity Scores within tolerance level (one-to-one)

$m(P_i) = min_j \{|P_i - P_j| < \epsilon\}$

- Step 1: Calculate the estimated propensity scores for treated and non-treated samples.
- Step 2: Matches for treated subject i are selected, if possible, only if $|P_i - P_j| < \epsilon$, where ϵ is a pre-specified tolerance.
- Step 3:The control subject with the value of P_j that is closest to P_i is selected as the match, and both are removed from the pool.
- Step 4: Repeat Step 2 and 3 until matching is no longer possible.
- Step 5: Discard the unmatched subjects.

Mahalanobis Metric Matching

- In this method, the subjects are ordered randomly, and then the distance between the first treated subject and all controls is calculated.
- This distance between treated subject i and untreated j (Mahalanobis distance):

$$d(i,j) = (u-v)^T C^{-1} (u-v)$$

where u and v are the matching variables for subject i and j, and C is the sample covariance matrix of the matching variables from the full set of control subjects.

- The control subject, j, with the minimum distance $d(i,j)$ is chosen as the match for treated subject i, and both subjects are removed from the pool.
- This process is repeated until matches are found for all treated subjects.

Mahalanobis Metric Matching

- In this method, the subjects are ordered randomly, and then the distance between the first treated subject and all controls is calculated.
- This distance between treated subject i and untreated j (Mahalanobis distance):

$$d(i,j) = (u-v)^T C^{-1} (u-v)$$

where u and v are the matching variables for subject i and j, and C is the sample covariance matrix of the matching variables from the full set of control subjects.

- The control subject, j, with the minimum distance $d(i,j)$ is chosen as the match for treated subject i, and both subjects are removed from the pool.
- This process is repeated until matches are found for all treated subjects.

Mahalanobis Metric Matching

- In this method, the subjects are ordered randomly, and then the distance between the first treated subject and all controls is calculated.
- This distance between treated subject i and untreated j (Mahalanobis distance):

$$d(i,j) = (u-v)^T C^{-1} (u-v)$$

  where u and v are the matching variables for subject i and j, and C is the sample covariance matrix of the matching variables from the full set of control subjects.
- The control subject, j, with the minimum distance $d(i,j)$ is chosen as the match for treated subject i, and both subjects are removed from the pool.
- This process is repeated until matches are found for all treated subjects.

Mahalanobis Metric Matching

- In this method, the subjects are ordered randomly, and then the distance between the first treated subject and all controls is calculated.
- This distance between treated subject i and untreated j (Mahalanobis distance):

$$d(i,j) = (u-v)^T C^{-1} (u-v)$$

  where u and v are the matching variables for subject i and j, and C is the sample covariance matrix of the matching variables from the full set of control subjects.
- The control subject, j, with the minimum distance $d(i,j)$ is chosen as the match for treated subject i, and both subjects are removed from the pool.
- This process is repeated until matches are found for all treated subjects.

Nearest available Mahalanobis metric matching within calipers defined by the propensity scores

- Step 1: Select a treated subject.
- Step 2: Choose all control subjects within a preset amount (or caliper) of the treated subject estimated logit of the propensity scores.
- Step 3: Include the logit of the estimated propensity scores with the other covariates in the calculation of the Mahalanobis distance.
- Step 4: Follow the procedure describe for Mahalanobis metric matching.
- Step 5: Discard all unmatched subjects.

Nearest available Mahalanobis metric matching within calipers defined by the propensity scores

- Step 1: Select a treated subject.
- Step 2: Choose all control subjects within a preset amount (or caliper) of the treated subject estimated logit of the propensity scores.
- Step 3: Include the logit of the estimated propensity scores with the other covariates in the calculation of the Mahalanobis distance.
- Step 4: Follow the procedure describe for Mahalanobis metric matching.
- Step 5: Discard all unmatched subjects.

Nearest available Mahalanobis metric matching within calipers defined by the propensity scores

- Step 1: Select a treated subject.
- Step 2: Choose all control subjects within a preset amount (or caliper) of the treated subject estimated logit of the propensity scores.
- Step 3: Include the logit of the estimated propensity scores with the other covariates in the calculation of the Mahalanobis distance.
- Step 4: Follow the procedure describe for Mahalanobis metric matching.
- Step 5: Discard all unmatched subjects.

Nearest available Mahalanobis metric matching within calipers defined by the propensity scores

- Step 1: Select a treated subject.
- Step 2: Choose all control subjects within a preset amount (or caliper) of the treated subject estimated logit of the propensity scores.
- Step 3: Include the logit of the estimated propensity scores with the other covariates in the calculation of the Mahalanobis distance.
- Step 4: Follow the procedure describe for Mahalanobis metric matching.
- Step 5: Discard all unmatched subjects.

Nearest available Mahalanobis metric matching within calipers defined by the propensity scores

- Step 1: Select a treated subject.
- Step 2: Choose all control subjects within a preset amount (or caliper) of the treated subject estimated logit of the propensity scores.
- Step 3: Include the logit of the estimated propensity scores with the other covariates in the calculation of the Mahalanobis distance.
- Step 4: Follow the procedure describe for Mahalanobis metric matching.
- Step 5: Discard all unmatched subjects.

Goal

Goal

We will use the previous techniques to match males and females based on physical charasteristics to remove potential bias due to size, and assess gender differences in pulmonary function and lung parenchyma on the matched subset of subjects.

Pulmonary functions measures (PFT) - Baseline Normal Characteristics by gender

(Before Matching)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
PFT measures:			
fev1/fvc	0.83	0.82	0.97
rv/tlc	0.29	0.34	0.004

(After Matching)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
PFT measures:			
rv/tlc	0.34	0.36	0.67

Pulmonary functions measures (PFT) - Baseline Normal Characteristics by gender

(Before Matching)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
PFT measures:			
fev1/fvc	0.83	0.82	0.97
rv/tlc	0.29	0.34	0.004

(After Matching)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
PFT measures:			
rv/tlc	0.34	0.36	0.67

Lung Parenchyma Measures (LP) - Baseline Normal Characteristics by gender

(Before Matching)(location = both)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
LP measures:			
mld	-862.87	-860.30	0.34
percent below -910	39.23	36.74	0.30
hu value 15 percentile	-935.01	-930.64	0.06
percent below -950	7.70	5.85	0.03
total volume	6914.10	5218.56	2.2×10^{-16}

(After Matching)(location = both)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
LP measures:			
total volume	6450.29	6043.71	0.23
hu value 15 percentile	-939.18	-929.69	0.02
percent below -950	8.84	4.64	0.01

Lung Parenchyma Measures (LP) - Baseline Normal Characteristics by gender

(Before Matching)(location = both)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
LP measures:			
mld	-862.87	-860.30	0.34
percent below -910	39.23	36.74	0.30
hu value 15 percentile	-935.01	-930.64	0.06
percent below -950	7.70	5.85	0.03
total volume	6914.10	5218.56	2.2×10^{-16}

(After Matching)(location = both)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
LP measures:			
total volume	6450.29	6043.71	0.23
hu value 15 percentile	-939.18	-929.69	0.02
percent below -950	8.84	4.64	0.01

Lung Parenchyma Measures (LP) - Baseline Normal Characteristics by gender

(Before Matching)(location = left)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
LP measures:			
mld	-862.98	-860.56	0.39
slope below -910	-1.63	-1.59	0.37
percent below -910	39.37	37.09	0.36
slope below -950	-2.31	-2.26	0.21
hu value 15 percentile	-935.50	-931.11	0.07
percent below -950	8.14	6.28	0.03
total volume	3272.10	2463.54	2.2×10^{-16}

(After Matching)(location = left)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
LP measures:			
total volume	2997.18	2824.22	0.23
hu value 15 percentile	-938.71	-929.97	0.04
percent below -950	8.94	4.89	0.01

Lung Parenchyma Measures (LP) - Baseline Normal Characteristics by gender

(Before Matching)(location = left)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
LP measures:			
mld	-862.98	-860.56	0.39
slope below -910	-1.63	-1.59	0.37
percent below -910	39.37	37.09	0.36
slope below -950	-2.31	-2.26	0.21
hu value 15 percentile	-935.50	-931.11	0.07
percent below -950	8.14	6.28	0.03
total volume	3272.10	2463.54	2.2×10^{-16}

(After Matching)(location = left)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
LP measures:			
total volume	2997.18	2824.22	0.23
hu value 15 percentile	-938.71	-929.97	0.04
percent below -950	8.94	4.89	0.01

Lung Parenchyma Measures (LP) - Baseline Normal Characteristics by gender

(Before Matching)(location = right)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
LP measures:			
mld	-868.68	-860.03	0.32
percent below -910	39.02	36.38	0.27
slope below -910	-1.65	-1.60	0.24
hu value 15 percentile	-934.44	-930.12	0.06
percent below -950	7.27	5.46	0.02
slope below -950	-2.47	-2.35	0.02
total volume	3642.00	2755.03	2.2×10^{-16}

(After Matching)(location = right)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
LP measures:			
total volume	3453.11	3219.49	0.27
hu value 15 percentile	-939.29	-929.42	0.02
percent below -950	8.68	4.43	0.01

Lung Parenchyma Measures (LP) - Baseline Normal Characteristics by gender

(Before Matching)(location = right)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
LP measures:			
mld	-868.68	-860.03	0.32
percent below -910	39.02	36.38	0.27
slope below -910	-1.65	-1.60	0.24
hu value 15 percentile	-934.44	-930.12	0.06
percent below -950	7.27	5.46	0.02
slope below -950	-2.47	-2.35	0.02
total volume	3642.00	2755.03	2.2×10^{-16}

(After Matching)(location = right)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
LP measures:			
total volume	3453.11	3219.49	0.27
hu value 15 percentile	-939.29	-929.42	0.02
percent below -950	8.68	4.43	0.01

Conclusion

- For pulmonary function measures before matching there was a statistically significant difference for rv/tlc. After matching the significance was completely removed. Therefore, rv/tlc is confounded by the gender size difference.
- Lung parenchyma measures of total volume before matching was statistically significant and after matching the significance was removed. Therefore, total volume is confounded by the gender size difference.
- For percent below $-950HU$ the difference was significant prior to matching and remained significant after matching. In this case, there might be a biological explanation.
- For HU value 15 percentile there was a marginal difference prior to matching and became statistically significant after matching. Therefore, before concluding anything we first must consider controlling demographics.

Conclusion

- For pulmonary function measures before matching there was a statistically significant difference for rv/tlc. After matching the significance was completely removed. Therefore, rv/tlc is confounded by the gender size difference.
- Lung parenchyma measures of total volume before matching was statistically significant and after matching the significance was removed. Therefore, total volume is confounded by the gender size difference.
- For percent below $-950HU$ the difference was significant prior to matching and remained significant after matching. In this case, there might be a biological explanation.
- For HU value 15 percentile there was a marginal difference prior to matching and became statistically significant after matching. Therefore, before concluding anything we first must consider controlling demographics.

Conclusion

- For pulmonary function measures before matching there was a statistically significant difference for rv/tlc. After matching the significance was completely removed. Therefore, rv/tlc is confounded by the gender size difference.
- Lung parenchyma measures of total volume before matching was statistically significant and after matching the significance was removed. Therefore, total volume is confounded by the gender size difference.
- For percent below $-950HU$ the difference was significant prior to matching and remained significant after matching. In this case, there might be a biological explanation.
- For HU value 15 percentile there was a marginal difference prior to matching and became statistically significant after matching. Therefore, before concluding anything we first must consider controlling demographics.

Conclusion

- For pulmonary function measures before matching there was a statistically significant difference for rv/tlc. After matching the significance was completely removed. Therefore, rv/tlc is confounded by the gender size difference.
- Lung parenchyma measures of total volume before matching was statistically significant and after matching the significance was removed. Therefore, total volume is confounded by the gender size difference.
- For percent below $-950HU$ the difference was significant prior to matching and remained significant after matching. In this case, there might be a biological explanation.
- For HU value 15 percentile there was a marginal difference prior to matching and became statistically significant after matching. Therefore, before concluding anything we first must consider controlling demographics.

Future work

- One to many or many to one matching.
- Explore best tolerance level for matching.
- Explore the noise induce by Mahalanobis metric matching within caliper defined by propensity scores.

Future work

- One to many or many to one matching.
- Explore best tolerance level for matching.
- Explore the noise induce by Mahalanobis metric matching within caliper defined by propensity scores.

Future work

- One to many or many to one matching.
- Explore best tolerance level for matching.
- Explore the noise induce by Mahalanobis metric matching within caliper defined by propensity scores.

References

Gomella LG, Haist SA(2011)

Chapter 18. Respiratory Care. *Clinician's Pocket Reference*

http://www.accesmedicine.com/content.aspx?aID=2695269

Rosenbaum, P.R. and Rubin, D.B.(1998)

Tutorial in Biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group.

Statistics in Medicine 17, 2265-2281.

Aknowledgements

Special thanks to:

Our mentor Dr. Gideon Zamba
ISIB Program
College of Public Health
Department of Biostatistics
University of Iowa, Iowa City

Partly supported by the University of Iowa Department of Biostatistics NHLBI-NCRR-NIH training grant (T15-HL097622-01 NHLBI).

APPENDIX

Full results tables

Pulmonary functions measures

Baseline Characteristic by gender
(Before Matching)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
bmi	26.48	24.22	< 0.01
Pulmonary Functions measures:			
fev1/fvc	0.83	0.82	0.97
fev1	1.09	1.081	0.80
fvc	1.04	1.04	0.79
tlc	1.03	1.04	0.60
svc	0.97	0.99	0.45
ic	1.24	1.19	0.27
dlco	1.13	1.19	0.27
rv/tlc	0.29	0.34	0.004

Pulmonary functions measures

Baseline Characteristic by gender
(After Matching by nearest logit PS within tolerance level)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
bmi	24.12	24.44	0.82
Pulmonary Functions measures:			
rv/tlc	0.34	0.36	0.67
tlc	1.09	1.12	0.67
fev1/fvc	0.81	0.84	0.27
fvc	0.98	1.05	0.20
ic	1.10	1.21	0.17
svc	0.90	0.98	0.17
dlco	0.97	1.13	0.10
fev1	0.99	1.09	0.09

Pulmonary Functions Measures

Baseline Characteristic by gender
(After Matching by nearest Mahalanobis metric with calipers defined by logit PS within tolerance level)

Variable	Male (N = 13) Mean	Female (N = 13) Mean	P-value
Demographics:			
age	32.92	29.31	0.55
height	1.74	1.74	0.83
weight	72.27	72.66	0.50
bmi	24.76	24.14	0.63
Pulmonary Functions measures:			
svc	0.97	0.97	0.97
ic	1.24	1.23	0.96
fvc	1.03	1.04	0.90
dlco	1.03	1.06	0.73
tlc	1.10	1.13	0.71
rv/tlc	0.31	0.34	0.52
fev1	1.04	1.08	0.51
fev1/fvc	0.81	0.85	0.06

Lung Parenchyma Measures

Baseline Characteristic by gender
(Before Matching)(location = both)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
bmi	26.48	24.22	¡0.01
Lung Parenchyma measures:			
mld	-862.87	-860.30	0.34
percent below -910	39.23	36.74	0.30
hu value 15 percentile	-935.01	-930.64	0.06
percent below -950	7.70	5.85	0.03
total volume	6914.10	5218.56	2.2×10^{-16}

Lung Parenchyma Measures

Baseline Characteristic by gender
(After Matching by nearest logit PS with tolerance level)(location = both)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
bmi	24.12	24.44	0.82
Lung Parenchyma measures:			
total volume	6450.29	6043.71	0.23
mld	-868.27	-859.48	0.07
percent below -910	43.77	35.24	0.06
hu value 15 percentile	-939.18	-929.69	0.02
percent below -950	8.84	4.64	0.01

Lung Parenchyma Measures

Baseline Characteristic by gender
(After Matching by nearest Mahalanobis metric with calipers defined by logit PS with tolerance level)(location = both)

Variable	Male (N = 13) Mean	Female (N = 13) Mean	P-value
Demographics:			
age	32.92	29.31	0.55
height	1.74	1.74	0.83
weight	75.27	72.66	0.50
bmi	24.76	24.14	0.63
Lung Parenchyma measures:			
total volume	6653.63	5958.65	0.04
mld	-870.55	-856.06	< 0.01
percent below -910	45.95	32.57	< 0.01
hu value 15 percentile	-941.31	-927.03	< 0.01
percent below -950	9.68	4.12	< 0.01

Lung Parenchyma Measures

Baseline Characteristic by gender
(Before Matching)(location = left)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
bmi	26.48	24.22	< 0.01
Lung Parenchyma measures:			
mld	-862.98	-860.56	0.39
slope below -910	-1.63	-1.59	0.37
percent below -910	39.37	37.09	0.36
slope below -950	-2.31	-2.26	0.21
hu value 15 percentile	-935.50	-931.11	0.07
percent below -950	8.14	6.28	0.03
total volume	3272.10	2463.54	2.2×10^{-16}

Lung Parenchyma Measures

Baseline Characteristic by gender
(After Matching by nearest logit PS within tolerance level)(location = left)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
bmi	24.12	24.44	0.82
Lung Parenchyma measures:			
slope below -910	-1.56	-1.57	0.87
slope below -950	-2.24	-2.27	0.71
total volume	2997.18	2824.22	0.23
mld	-866.68	-859.46	0.15
percent below -910	42.54	35.33	0.12
hu value 15 percentile	-938.71	-929.97	0.04
percent below -950	8.94	4.89	0.01

Lung Parenchyma Measures

Baseline Characteristic by gender
(After Matching by nearest Mahalanobis metric with calipers defined by logit PS within tolerance level)(location = left)

Variable	Male (N = 13) Mean	Female (N = 13) Mean	P-value
Demographics:			
age	32.92	29.31	0.55
height	1.74	1.74	0.83
weight	75.27	72.66	0.50
bmi	24.76	24.14	0.63
Lung Parenchyma measures:			
slope below -950	-2.25	-2.25	0.93
slope below -910	-1.52	-1.60	0.21
total volume	3122.72	2778.19	0.03
percent below -910	45.22	32.59	0.01
mld	-869.48	-855.90	< 0.01
hu value 15 percentile	-941.24	-927.22	< 0.01
percent below -950	9.90	4.36	< 0.01

Lung Parenchyma Measures

Baseline Characteristic by gender
(Before Matching)(location = right)

Variable	Male (N = 51) Mean	Female (N = 74) Mean	P-value
Demographics:			
age	32.78	35.01	0.41
height	1.79	1.65	2.2×10^{-16}
weight	85.24	65.83	6.1×10^{-16}
bmi	26.48	24.22	0.00
Lung Parenchyma measures:			
mld	-868.68	-860.03	0.32
percent below -910	39.02	36.38	0.27
slope below -910	-1.65	-1.60	0.24
hu value 15 percentile	-934.44	-930.12	0.06
percent below -950	7.27	5.46	0.02
slope below -950	-2.47	-2.35	0.02
total volume	3642.00	2755.03	2.2×10^{-16}

Lung Parenchyma Measures

Baseline Characteristic by gender
(After Matching by nearest logit PS within tolerance level)(location = right)

Variable	Male (N = 11) Mean	Female (N = 11) Mean	P-value
Demographics:			
age	33.36	31.00	0.74
height	1.74	1.74	0.87
weight	73.39	73.67	0.95
bmi	24.12	24.44	0.82
Lung Parenchyma measures:			
slope below -910	-1.57	-1.59	0.75
slope below -950	-2.39	-2.32	0.60
total volume	3453.11	3219.49	0.27
mld	-869.45	-859.49	0.04
percent below -910	44.67	35.14	0.03
hu value 15 percentile	-939.29	-929.42	0.02
percent below -950	8.68	4.43	0.01

Lung Parenchyma Measures

Baseline Characteristic by gender
(After Matching by nearest Mahalanobis metric with calipers defined by logit PS within tolerance level)(location = right)

Variable	Male (N = 13) Mean	Female (N = 13) Mean	P-value
Demographics:			
age	32.92	29.31	0.55
height	1.74	1.74	0.83
weight	75.27	72.66	0.50
bmi	24.76	24.14	0.63
Lung Parenchyma measures:			
slope below -950	-2.36	-2.32	0.72
slope below -910	-1.54	-1.62	0.19
total volume	3530.92	3180.46	0.07
mld	-871.28	-855.17	< 0.01
percent below -910	46.43	32.52	< 0.01
percent below -950	9.42	3.91	< 0.01
hu value 15 percentile	-941.09	-926.77	< 0.01