

Correlation between Primary Tumor Location and Age in nonsmokers

BSMS222 Jihyun Seo(2015250126)

1. Introduction

In resource paper 'Proteogenomics of Non-smoking Lung Cancer in East Asia Delineates Molecular Signature of Pathogenesis and Progression', focusing on clinical data from Taiwan(TW) cohort, what I was wondering is about primary tumor location in treatment-naive patients from TW, especially whose histology type is ADC(adenocacinoma) and who are nonsmokers. Which factor among paient's characters such as age, gender, ADC stage or EGFR_status would have correlation to tumor location?

Description of tumor location in lung. RUL = Right Upper Lobe, RML = Right Middle Lobe, RLL = Right Lower Lobe LUL = Left Upper Lobe, LLL = Left Lower Lobe

2. Dataset & Visualizing

```
library(dplyr)
```

```
##  
## 다음의 패키지를 부착합니다: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##   filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
library(ggplot2)  
library(readxl)  
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v tibble  3.1.4    v purrr   0.3.4  
## v tidyr   1.1.3    v stringr 1.4.0  
## v readr   2.0.1    v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()    masks stats::lag()
```

```
d1 <- read_excel("mmc1.xlsx", sheet = "Table S1A_clinical_103patient")
head(d1)
```

ID	Proteome_Batch	Gen...	Age	Smoking Status	Histology Type	St...	EGFR_Stat...
<chr><chr>		<chr>	<dbl>	<chr>	<chr>	<chr>	<chr>
P002 B01-2		Male	73.77687	Nonsmoke	ADC	IB	others
P004 B01-4		Female	52.97741	Nonsmoke	SCC	IA	exon19del
P005 B02-1		Male	72.75017	Current_Smoker	SCC	IA	WT
P006 B02-2		Female	46.86105	Nonsmoke	ADC	IB	WT
P007 B02-3		Male	67.40589	Nonsmoke	ADC	IIA	WT
P009 B03-1		Female	53.80424	Nonsmoke	ADC	IIA	L858R

6 rows | 1-8 of 9 columns

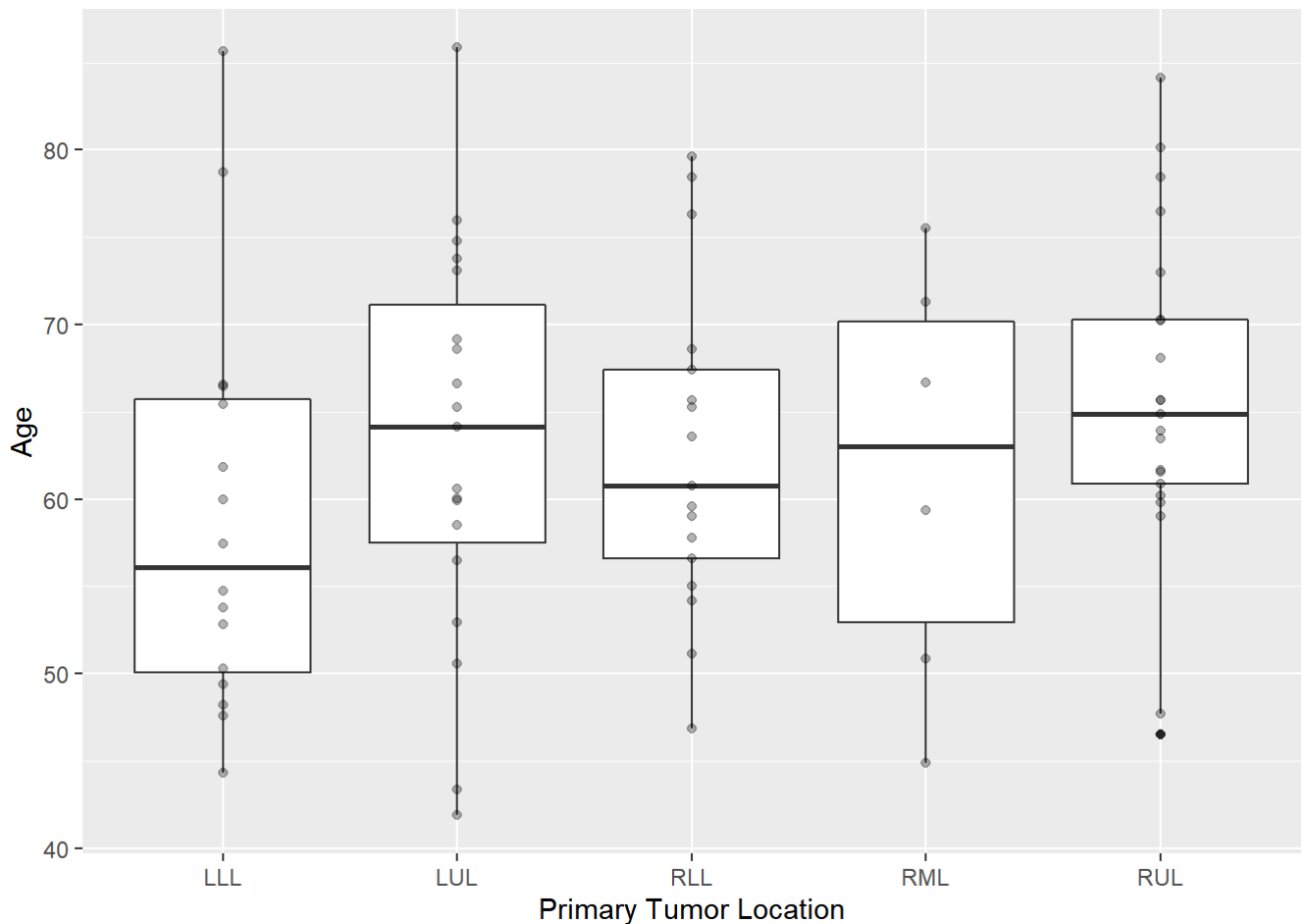
First of all, select data of only nonsmokers and ADC patients from the clinical data source given from paper.

```
d2 <- d1 %>% filter(`Smoking Status` == 'Nonsmoke', `Histology Type` == "ADC")
summary(d2)
```

```
##      ID      Proteome_Batch      Gender      Age
## Length:79      Length:79      Length:79      Min.   :41.88
## Class :character Class :character Class :character 1st Qu.:55.74
## Mode  :character Mode  :character Mode  :character Median :61.86
##                                     Mean  :62.80
##                                     3rd Qu.:68.88
##                                     Max.   :85.86
## Smoking Status Histology Type      Stage      EGFR_Status
## Length:79      Length:79      Length:79      Length:79
## Class :character Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character Mode  :character
##
##
## Primary Tumor Location
## Length:79
## Class :character
## Mode  :character
##
##
##
```

Four points of view 1) gender and tumor location 2) stage and tumor location 3) EGFR_Status and tumor location 4) age and tumor location Among these, 4)age and tumor location had meaningful correlation.

```
d2 %>% ggplot(aes(`Primary Tumor Location`, Age)) +
  geom_boxplot() +
  geom_point(alpha=0.3, size = 1.5)
```



3. Discussion

According to the figure, in left lung lobes, median of upper lobe was higher in age than one of lower lobe. Likewise in right lung lobes, median of upper part was the highest, followed by middle part and then lower part. In terms of median, left and right lungs both look right-top direction. And it leads to positive correlation between age and location of ADC(upper-middle-lower).

4. Reference

Yi-JuChen, Theodoros I.Roumeliotis, Ya-Hsuan Chang, Ching-Tai Chen, Chia-Li Han, Miao-Hsia Lin, Huei-Wen Chen , Gee-Chen Chang, Yih-Leong Chang, Chen-Tu Wu, Mong-Wei Lin, Min-Shu Hsieh, Yu-Tai Wang, Yet-Ran Chen, Inge Jonassen, Fatemeh Zamanzad Ghavidel, Ze-Shiang Lin, Kuen-Tyng Lin1 ...Yu-Ju Chen, Proteogenomics of Non-smoking Lung Cancer in East Asia Delineates Molecular Signatures of Pathogenesis and Progression, Cell, Volume 182, Issue 1, 9 July 2020, Pages 226-244.e17