Module_2: *Fibrosis*

Team Members:

Addison and Mel

Project Title:

Predicting the extent of lung fibrosis at varying lung depths

Project Goal:

This project seeks to answer… *How does the extent of fibrosis vary throughout the depth of a fibrotic lung? And, using interpolation, can you predict the amount of fibrosis at a specific depth into the lung?*

Disease Background:

- Prevalence & incidence *Though lung fibrosis is a rare disease, both its prevalence and incidence are increasing rapidly. Specifically, its prevalence has been increasing since the COVID-19 pandemic. In North America the prevalence rates are between 2.4-2.98 for every 10,000 people, however rates are higher in other parts of the world such as South Korea. (https://pmc.ncbi.nlm.nih.gov/articles/PMC9872080/). Another study states that the prevalence is about 27.2 per 100,000 people in North America, while the incidence is about 9.0 per 100,000 people in North America. (https://pmc.ncbi.nlm.nih.gov/articles/PMC12330001/). While this does still show that lung fibrosis is a rare disease, its increasing rates are a cause for concern.*
- Risk factors (genetic, lifestyle) *Some lifestyle risk factors are smoking, certain types of work, and cancer treatments. Specifically, those that are exposed to toxins or pollutants such as silica dust, asbestos, hard metal dusts, mold, or wood, coal and grain dusts frequently may cause lung fibrosis. In terms of cancer treatment, their effects are seen through the impacts of radiation treatments, which may damage the lungs. Furthermore, the types of medicine that a person takes may predispose them to lung damage. Heart medicines, antibiotics, and anti-inflammatory medicines may cause lung damage, in addition to certain medical conditions such as lupus, pneumonia, and dermatomyositis. (https://www.mayoclinic.org/diseases-conditions/pulmonary-fibrosis/symptoms-causes/syc-20353690)*
- Symptoms *Symptoms of lung fibrosis are shortness of breath, dry cough, extreme tiredness, unintended weight loss, muscle and joint aches, and clubbing (widening and rounding) or the fingertips or toes. The progression of these symptoms and their severity, as some cases exhibit very quick and severe progression, while other cases exhibit moderate and slow progression. If a patient with lung fibrosis suddenly sees a worsening in shortness of breath, this is considered an extreme case of acute exacerbation. (https://www.mayoclinic.org/diseases-conditions/pulmonary-fibrosis/symptoms-causes/syc-20353690)*
- Standard of care treatment(s) *There are several standard of care treatments based on the severity and progression of lung fibrosis. For less severe cases, a patient may receive*

pulmonary rehabilitation with exercise, information on how to make lifestyle changes, and other intervention methods in order to control the progression of the disease. A person may also receive medication in the form of immunosuppressants and steroids. Because lung fibrosis is caused by scar tissue, there are also two anti-scarring methods available in order to treat IPF. For more severe cases, a patient may receive supplemental oxygen through a tube or mask, or will receive a lung transplant in the most severe cases. (https://www.dukehealth.org/treatments/lung-disease/pulmonary-fibrosis-and-interstitial-lung-disease)

- Biological mechanisms (anatomy, organ physiology, cell & molecular physiology) *Within the lungs, fibrosis affects the thin connective tissue between lung alveoli and capillaries, specifically the alveolar walls and interstitium. As tissue becomes fibrotic, this thin tissue scars, which reduces the surface area of healthy lungs. Additionally, lung elasticity decreases, lung diffusion capacity (how well oxygen and carbon dioxide move between blood and the lungs) decreases, and lung alveoli stiffen. All of these factors cause strain on the lungs and can make it more difficult to breathe. Typically, the disease starts in the lower and peripheral lungs as dysregulation in wound healing, consistent with the scarring found in lung fibrosis. In this process, damaged alveolar cells cause fibroblasts to turn into myofibroblasts and overproduce collagen and extracellular matrix proteins, stiffening the tissue. This process of fibroblasts turning to myofibroblasts is driven by the release of cytokines and growth factors. Myofibroblasts allow scar tissue to form as there is abnormal apoptosis (process of cell death) control, and the result is irreversible cell damage. (ChatGPT prompt: Summarize the anatomy, organ physiology, cell & molecular physiology of lung fibrosis)*

Data-Set:

- Who and where was the data obtained? What is being analyzed? *The data was obtained from the Peirce-Cottler Lab at UVA. It contains 78 black and white .jpg scans of the layers of a single mouse lung. The black pixels represent healthy lung tissue, while the white pixels represent fibrotic lesions. The data analyzed will be the percent of white pixels in comparison to the percent of black pixels to represent the ratio of fibrotic lesions to healthy tissue. An important note: blood vessels are also stained and shown in large, white blots, resulting in higher error for the amount of lesions in tissue.*
- How was the data obtained and what techniques were used? *The lab used a Bleomycin-induced Lung Injury Model, ultimately using the drug Bleomycin to stimulate lung fibrosis in the mice. The mice then have their lungs humanely harvested, which are then fixed in paraformaldehyde and gel/wax, transversely and thinly sliced, and stained on a glass slide. The stain binds to and colors the desired proteins and each slice is then scanned and digitally uploaded. Using image editing software, the color corresponding to myofibroblasts, the cells contributing to the fibrotic scarring, is finally isolated and made black and white.*
- Sources: *Hannan, R. T., Miller, A. E., Hung, R. C., Sano, C., Peirce, S. M., & Barker, T. H. (2020). Extracellular matrix remodeling associated with bleomycin-induced lung injury supports pericyte-to-myofibroblast transition. Matrix biology plus, 10, 100056. https://doi.org/10.1016/j.mbplus.2020.100056 ; https://canvas.its.virginia.edu/courses/153653/files/16745297?module_item_id=1821588*

Data Analyis:

Import packages, Import images, Write depths, Make csv

```
'''Module 2: count black and white pixels in a .jpg and extrapolate
points'''
# Importing necessary modules

from termcolor import colored
import cv2
import numpy as np
import matplotlib.pyplot as plt
from scipy.interpolate import interp1d
import pandas as pd

# Load the images you want to analyze

filenames = [
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010017.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010039.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010061.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010071.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010164.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010066.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010087.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010089.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010093.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010096.jpg",
r"/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010030.jpg"
]

# Enter the depth of each image (in the same order that the images are
listed above; you can find these in the .csv file provided to you
which is tilted: "Filenames and Depths for Students")

depths = [
45,
15,
585,
```

```
7100,
2200,
1000,
8000,
10000,
9300,
9400,
200
]

# Make the lists that will be used

images = []
white_counts = []
black_counts = []
white_percents = []

# Build the list of all the images you are analyzing

for filename in filenames:
    img = cv2.imread(filename, 0)
    images.append(img)

# For each image (until the end of the list of images), calculate the
number of black and white pixels and make a list that contains this
information for each filename.

for x in range(len(filenames)):
    _, binary = cv2.threshold(images[x], 127, 255, cv2.THRESH_BINARY)

    white = np.sum(binary == 255)
    black = np.sum(binary == 0)

    white_counts.append(white)
    black_counts.append(black)

# Print the number of white and black pixels in each image.

print(colored("Counts of pixel by color in each image", "yellow"))
for x in range(len(filenames)):
    print(colored(f"White pixels in image {x}: {white_counts[x]}",
"white"))
    print(colored(f"Black pixels in image {x}: {black_counts[x]}",
"black"))
    print()

# Calculate the percentage of pixels in each image that are white and
make a list that contains these percentages for each filenamae

for x in range(len(filenames)):
```

```
    white_percent = (100 * (white_counts[x] / (black_counts[x] +
white_counts[x])))
    white_percents.append(white_percent)

# Print the filename (on one line in red font), and below that line
print the percent white pixels and depth

print(colored("Percent white px:", "yellow"))
for x in range(len(filenames)):
    print(colored(f'{filenames[x]}:',"red"))
    print(f'{white_percents[x]}% White | Depth: {depths[x]} microns')
    print()

'''Write your data to a .csv file'''

# Create a DataFrame that includes the filenames, depths, and
percentage of white pixels
df = pd.DataFrame({
    'Filenames': filenames,
    'Depths': depths,
    'White percents': white_percents
})

# Write that DataFrame to a .csv file

df.to_csv('Percent_White_Pixels.csv', index=False)

print("CSV file 'Percent_White_Pixels.csv' has been created.")

'''the .csv writing subroutine ends here'''

# Interpolate a point: given a depth, find the corresponding white
pixel percentage

interpolate_depth = float(input(colored("Enter the depth at which you
want to interpolate a point: ", "yellow")))

x = depths
y = white_percents

i = interp1d(x, y, kind='cubic')  # You can also use 'quadratic',
'cubic', etc.
interpolate_point = i(interpolate_depth)
print(colored(f'The interpolated point is at the x-coordinate
{interpolate_depth} and y-coordinate {interpolate_point}.', "green"))

depths_i = depths[:]
depths_i.append(interpolate_depth)
white_percents_i = white_percents[:]
white_percents_i.append(interpolate_point)
```

```
# Make two plots: one that doesn't contain the interpolated point,
just the data calculated from your images, and one that also contains
the interpolated point (shown in red)
fig, axs = plt.subplots(2, 1)

axs[0].scatter(depths, white_percents, marker='o', linestyle='-',
color='blue')
axs[0].set_title('Plot of depth of image vs percentage white pixels')
axs[0].set_xlabel('depth of image')
axs[0].set_ylabel('white pixels as a percentage of total pixels')
axs[0].grid(True)


axs[1].scatter(depths_i, white_percents_i, marker='o', linestyle='-',
color='blue')
axs[1].set_title('Plot of depth of image vs percentage white pixels w/
interpolated point (red)')
axs[1].set_xlabel('depth of image')
axs[1].set_ylabel('white pixels as a percentage of total pixels')
axs[1].grid(True)
axs[1].scatter(depths_i[len(depths_i)-1],
white_percents_i[len(white_percents_i)-1], color='red', s=100,
label='Highlighted point')


# Adjust layout to prevent overlap
plt.tight_layout()
plt.show()
```

```
Counts of pixel by color in each image
White pixels in image 0: 27561
Black pixels in image 0: 4166743

White pixels in image 1: 21648
Black pixels in image 1: 4172656

White pixels in image 2: 56003
Black pixels in image 2: 4138301

White pixels in image 3: 117756
Black pixels in image 3: 4076548

White pixels in image 4: 65021
Black pixels in image 4: 4129283

White pixels in image 5: 60715
Black pixels in image 5: 4133589

White pixels in image 6: 155019
Black pixels in image 6: 4039285
```

```
White pixels in image 7: 369855
Black pixels in image 7: 3824449

White pixels in image 8: 168932
Black pixels in image 8: 4025372

White pixels in image 9: 179894
Black pixels in image 9: 4014410

White pixels in image 10: 37799
Black pixels in image 10: 4156505

Percent white px:
/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010017.jpg:
0.6571054458618164% White | Depth: 45 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010039.jpg:
0.5161285400390625% White | Depth: 15 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010061.jpg:
1.3352155685424805% White | Depth: 585 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010071.jpg:
2.8075218200683594% White | Depth: 7100 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010164.jpg:
1.5502214431762695% White | Depth: 2200 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010066.jpg:
1.4475584030151367% White | Depth: 1000 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010087.jpg:
3.6959409713745117% White | Depth: 8000 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010089.jpg:
8.81803035736084% White | Depth: 10000 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
Slobe ch010093.jpg:
4.027652740478516% White | Depth: 9300 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_SK658
```

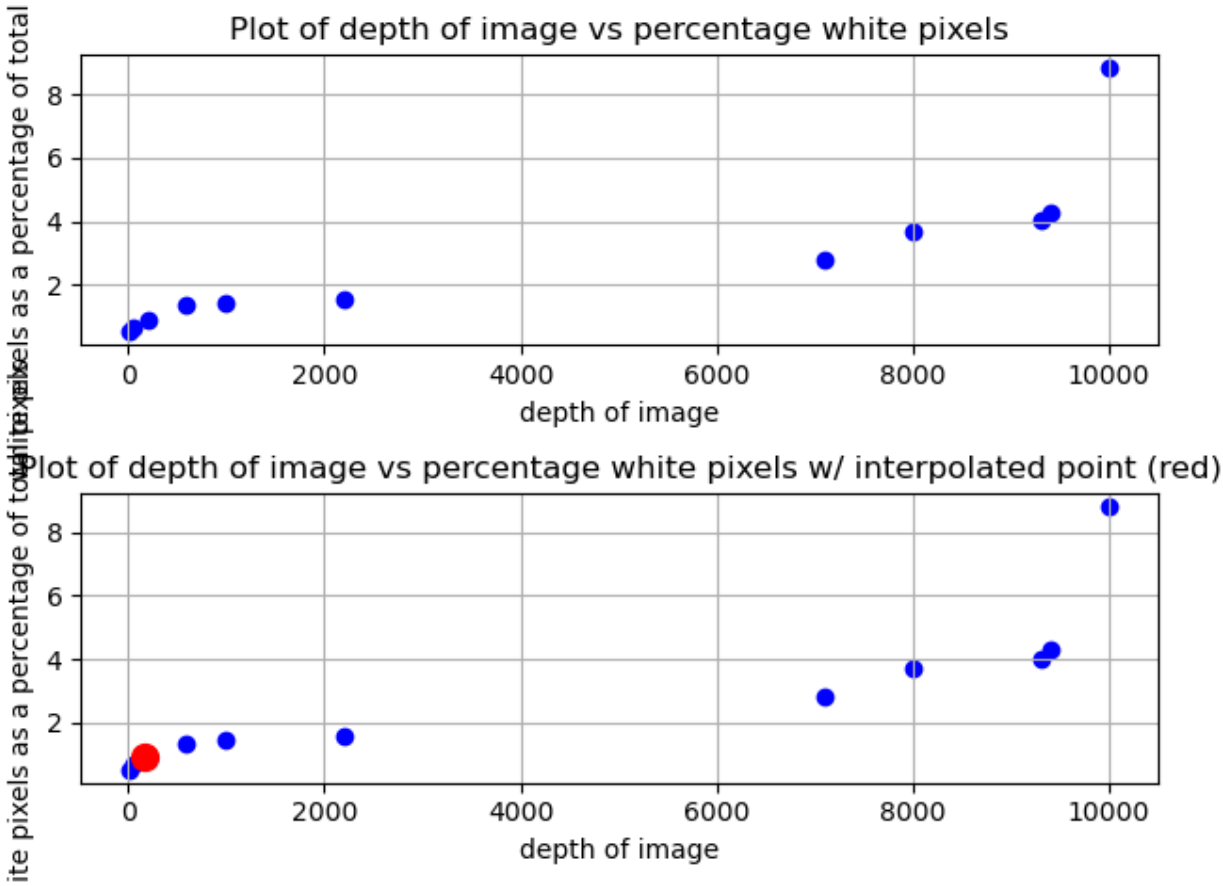
```
Slobe ch010096.jpg:
4.289007186889648% White | Depth: 9400 microns

/Users/Melanie/Downloads/Comp_BME/Module2_LungFibrosis/MASK_Sk658
Llobe ch010030.jpg:
0.9011983871459961% White | Depth: 200 microns

CSV file 'Percent_White_Pixels.csv' has been created.

Enter the depth at which you want to interpolate a point:  175

The interpolated point is at the x-coordinate 175.0 and y-coordinate
0.8861627646202846.
```



Verify and validate your analysis:

Verification:

- Does your plotted "percent white pixels" data support that fibrosis increases as you look deeper into the lung? *For lower depths, yes, the fibrosis increases as you look deeper into the lung. When analyzing depths from 0-7100, the scatterplots showed that there was rapid increase at lower depths and more gradual increase at higher depths. However, from depths 7100-10000, a slightly different trend was observed. While there*

was a general increase over this range, there was a slight dip in the percent of white pixels around 8000-9300. See the graphs generated in data analysis for more information. Overall though, the data shows an increase in percentage of white pixels as you travel deeper into the lung.

- Do your estimated values make sense compared to the other data points you measured? *When estimating the value for percent white pixels at a depth of 8100 microns, a y-value of 3.511524 was interpolated using linear methods. When estimating the value for percent white pixels at a depth of 8100 microns, a y-value of 3.513928 was interpolated using quadratic methods. When estimating the value for percent white pixels at a depth of 8100 microns, a y-value of 3.499590 was interpolated using cubic methods. These values are only a few decimal places away from one another, so the three interpolations at this value do not have significant differences. When compared to data in a similar range of 8100 microns, the percent values are at around a high 3, low 4 percent. These models are thus in line with our data set with six images. With another six images and interpolating the data at a shallower depth of 175 microns, the program yielded a percent of white pixels of 0.820354 at a linear method, 1.102377 for a quadratic method, and 1.087061 for a cubic method. The percentages between a quadratic and cubic model are again similar, while the linear model deviates from the two models by a few tenth places. The percent of white pixels at a close value to 175 microns, 200, yielded 0.901198. Thus, the method closest to the actual data is the linear method and can be named our most accurate method out of the three, at least for interpolations on the lower end of the range. The aforementioned interpolations were made with six images in the data set. Increasing the data to 11 images resulted in the following predicted percent of white pixels at 175 microns: 0.861828 for a linear method, 0.879761 for quadratic, and 0.886162 for cubic. These results differ from using solely six images since now the values are much closer together, varying by a few hundredths of decimal places. All three values are still close to the actual percent value for 200 microns though the cubic function is closest, therefore making a cubic interpolation the most accurate method for this larger data set.*
- Do your results align with published literature on the topic? *Our results are supported by published findings which concluded that the extent of fibrosis is associated with lung morphology (https://pmc.ncbi.nlm.nih.gov/articles/PMC9796832/). While our findings were specifically based on lung depth, various changes in lung structures may impact the extent of fibrosis at various depths. The publication may also support conclusions that the "dip" found in fibrotic tissue at around 8000-9300 microns may be due to changes in lung structure.*

Validation

- How does the percent white pixels that you measure in the real image compare to your interpolated value at/around that same depth? *One image was found at a depth of 8100 microns, and the code to find its percent white pixels is shown in the data analysis section. The percent found was 3.2827377319335938%. This value is slightly lower than what was interpolated using linear, quadratic, and cubic models (3.511524, 3.513928, and 3.499590, respectively). However, it does make sense that the actual value would be lower because when using models of higher degrees, the value of percent white pixels at 8100 microns decreased. It would be reasonable to assume that as the degree of models continued to increase, the percent of white pixels at 8100 microns would continue to approach the actual value.*

Conclusions and Ethical Implications:

Generally, the overall trend shows that fibrosis in this mouse model increases with lung depth, particularly in the range of 0-7100 microns. There is a slight dip in the range of 8000-9300 microns, however this may not be a characteristic of lung fibrosis and may instead be due to imaging issues or changes in lung structure. It was also found that the various methods of interpolation were generally consistent, however the cubic model was the most accurate with the actual white pixel count. This shows that interpolation models of a higher degree may be more accurate in their estimates.

In the creation of a new lung biopsy device, it is recommended that the biopsy device is capable of taking samples at lower lung depths. This is because the amount of fibrotic tissue increases with lung depth. Additionally, the biopsy device should be able to take samples from very specific lung depths in order to avoid any areas or lung structures where there may be a "dip" in fibrosis. This modification would avoid any underestimates of the extent of fibrosis in a patient.

A new device supported by these results would likely improve diagnostic methods of lung fibrosis, increasing patient care. However, it is also important that a new device does not solely rely on mouse model data, results should be validated through human biopsy in order to ensure that the device is built off of the most accurate results and that the trend found in the mouse model is consistent in a human model. Additionally, our work necessitates accuracy and clear communication of results in order to not misinform the design of the new medical device and to avoid harm towards patients. Lastly, the results should be used to improve diagnoses across all patient groups, no matter the socioeconomic background/insurance coverage.

Limitations and Future Work:

- Error of excess white pixels as a result of blood vessel stains: Due to the fact that blood vessel stains from the images also appear white in the images, the calculated percentages for those images containing blood vessels will be higher than the actual values. This could also be one explanation for why the interpolated values for one depth were higher than what was expected. In order to work around this limitation, the code would have to be able to determine the difference between fibrosis pixels and blood vessel pixels in the image. This may be done by filtering out large groups of white pixels as the area of blood vessels appearing in the image is normally larger than that of fibrosis. A machine learning program may be trained to recognize these differences in the future.
- Mouse lungs have different biology than human lungs, although they are both still relatively similar. While it may be necessary to use human biopsies of the lung in order to corroborate more accurate results, the ethical implications of that study must be considered as this number of biopsies cannot be performed on a living patient. Other non-invasive methods of study may be used to determine the extent of fibrosis in human lung tissue if they become available and accessible. Otherwise, future data collection and analysis should be made with human lungs voluntarily donated from patients once they are deceased. Patients must be informed of where their organs will go and for what purpose, and they must be in able-bodied condition to explicitly confirm consent without coercion. These measures must be taken to maintain ethics in lung fibrosis studies.

NOTES FROM YOUR TEAM:

- Class 10/2: Met team, learned about the project and data set, and started initial data analysis
- Class 10/7: Showed csv files, learned about interpolation, discussed project work breakdown for learning about the disease and learning about the data set.
- Class 10/9: Consolidated both of our analysis code into one Jupyter notebook, began validation and verification of our data, reflected on conclusions, limitations, and future work
- Week of 10/13: Finished remaining parts of the notebook, including background information, validation, verification, conclusion, limitations, and future work, copied final code from .py to Jupyter notebook, finalized Jupyter notebook for submission

QUESTIONS FOR YOUR TA:

N/A