Report from 6/14:

One group (David E., Lindsey, Onyekachi, Isaac) looked at the summary tables in some of the literature to create probabilities of a particular feature occurring in a severe case. This was done for the Covid and flu literature (see CovidFlu.xls) and lines 19, 22, and 45 in the spreadsheet regarding COPD (see COPD.xls). Work is also being done to also get the mean and standard deviation for certain characteristics as well. Repeated variables have been averaged by weight.

Once those numbers have been finalized, then Hanna has written a code that will take those probabilities and create synthetic patient data. This will then provide a database which the models can test against.

Another group (Scott, Konda, Victoria, Khoa, Yixuan, Elisabeth) is working on a multivariate probability model. The plan for this is to evaluate the current distributions in use and look into dependence and correlation between symptoms. We are planning to investigate the use of conditional probability influenced distributions, truncated distributions, and eventually into multivariate distributions to generate more realistic patients. Current outline of steps: 1. gather the current distributions and variables in use. 2. Separate them into ones that should be truncated or dependent on each other (using correlations analysis or independence testing). 3. Alter the distributions. 4. Test how these distributions work in generating more realistic people. 5. Big Goal: Begin to develop an overall multivariate distribution. Elisabeth thought of the possibility of adding a coefficient or parameter that captures risk behavior of the population.

Another group doing literature review (Hunter, Wei doing general COPD review), Wyatt is doing allergens.

Wyatt: Looking specifically at the relationship between mold as an allergen and its effects on bronchiectasis in COPD. I put a paper on in it in the literature subgroup files. It appears that there is a link to an increased response to patients that have COPD and pack years of smoking versus patients that solely have a risk factor of pack years of smoking.

Hunter: I looked into genetic factors that could affect COPD exacerbation. My thought process is that genotyping is so cheap and easy now, it is feasible to get that when someone subscribes to Vironix. According to GOLD, the only known significant genetic risk factor is alpha-1 antitrypsin. I may have just been down a bad rabbit hole. I also found a paper that was predicting outcomes of patients with COPD exacerbations that were at the ER

ER Paper: <https://erj.ersjournals.com/content/erj/32/4/953.full.pdf>

Rabbit hole: <https://pubmed.ncbi.nlm.nih.gov/9246142/>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1746834/>

Wei:

1. Learning more general information about COPD: <https://www.healthline.com/health/copd>

1. Four GOLD grades based on spirometry testing: (<https://goldcopd.org/2021-gold-reports/>)

1. Grade 1: mild

2. Grade 2: moderate

3. Grade 3: severe

4. Grade 4: very severe

Question: The existing data covers grades 2-4, shall we find the data including grade 1 as well? In this grade, a person may get short of breath/cough several days a week. The following website mentioned that “You have no more than one exacerbation per year of your COPD and you are not hospitalized for it.” However, considering the development of the disease, it might be good to include the grade 1 data to the analysis.

1. More details about the descriptions of different stages/groups:

<https://www.healthline.com/health/copd/stages#groups>

1. In the COPD.xlsx file’s first sheet BMC, the data are in a percentage scale from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1764756/>
2. A possible good reference:

<https://respiratory-research.biomedcentral.com/articles/10.1186/1465-9921-11-122>

Reason: This paper seems focus on the smoking factor and different four GOLD stages.

**Whiteboards**

Monday: