Exploring Longitudinal Pulmonary Exacerbation Outcome Trajectories in Cystic Fibrosis Patients

Brenton Graham

12/5/22

Cystic Fibrosis (CF)

- Hereditary disease which involves production of abnormally viscous mucus
- Early airway vulnerability to chronic bacterial infections
- Most common life-threatening autosomal recessive disease in US
  - Affects 1 in 4000 newborns
  - No known cure

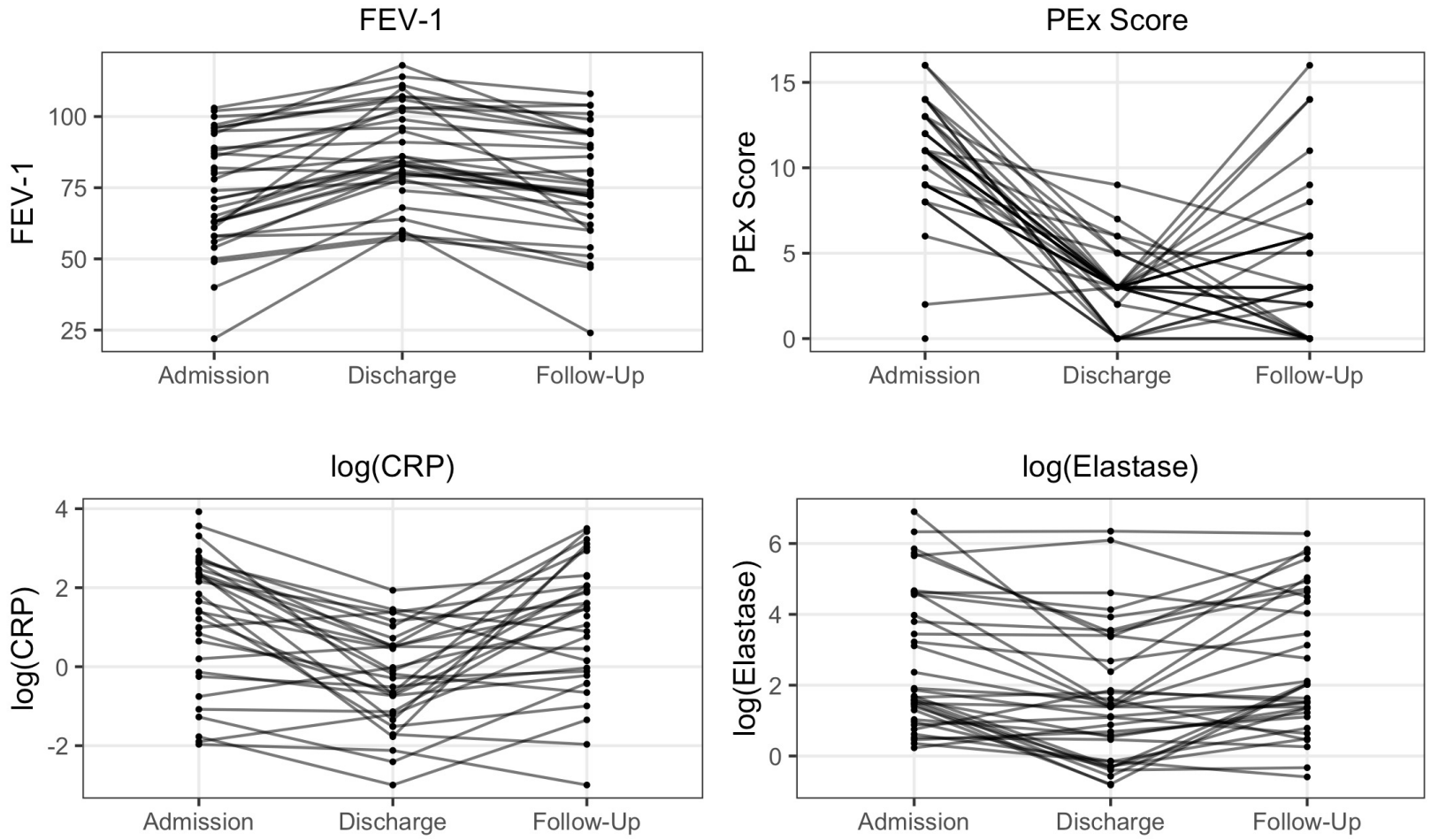
Pulmonary Exacerbations (PExs)

- Significant life events (“lung attacks”) in people with CF
- Associated with declining lung function, reduced quality of life, hospitalizations, and decreased survival
- A leading cause of morbidity in CF
- Treatment response often suboptimal despite seemingly appropriate antimicrobial therapy
- PEx-related outcomes (markers)
  - FEV-1: Measure of lung function
  - PEx Score: Self-evaluated score of PEx severity
  - CRP: Marker of inflammation
  - Neutrophil Elastase: Secreted during inflammation; destroys bacteria and host tissue

Study Design

- Participants include CF patients hospitalized at the onset of a PEx
  - 34 CF patients and 39 PExs (5 subjects with multiple PEx)
- Subjects evaluated at three timepoints separated by 10-14 days
  - **T1:** Admission (PEx onset) – ***least healthy***
  - **T2:** Hospital discharge – ***most healthy***
  - **T3:** Follow-up – ***somewhere in between***
- At each time point subjects were evaluated for…
  - PEx-related clinical outcomes
  - CF pathogen detection (Culture and 16S)

PEx-Related Outcomes



Project Aim

- **Outcome of Interest:** Neutrophil elastase
  - Inflammatory marker that suggests lung disease
  - “If you can detect elastase, disease is in progression”
- **Question:** Are there any baseline risk factors associated with outcome trajectory (treatment response)?
- **Hypotheses**
  - Hypothesis 1: There exist sub-groups in our population that experience separable outcome trajectories (treatment response)
  - Hypothesis 2: There exist baseline risk factors that are associated with said sub-groups

Latent Class Analysis (LCA)

- LCA is a longitudinal (LMM) framework that can be used for **clustering** repeated responses based on similar patterns
  - This project: we want to cluster subjects based on treatment response over time
- Longitudinal piece
  - LMM framework
  - Can specify random effects and correlation structures that shape
- Fitting a latent class mixed model
  - lcmm package in R
  - Must specify the number of latent classes
  - Model selection process used to determine the optimal # of classes

Latent Class Mixed Model

- Model specifications
  - Fixed effects: Time (as class variable)
  - Random effects
    - Random intercept for subject
    - Random effect for time
  - Correlation structure
    - Tried both unstructured and UN(1) equivalent from SAS
  - Number of latent classes
    - Varied between 2-5 classes

Unstructured UN

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} & \sigma_{42} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix}$$

Banded main diagonal UN(1)

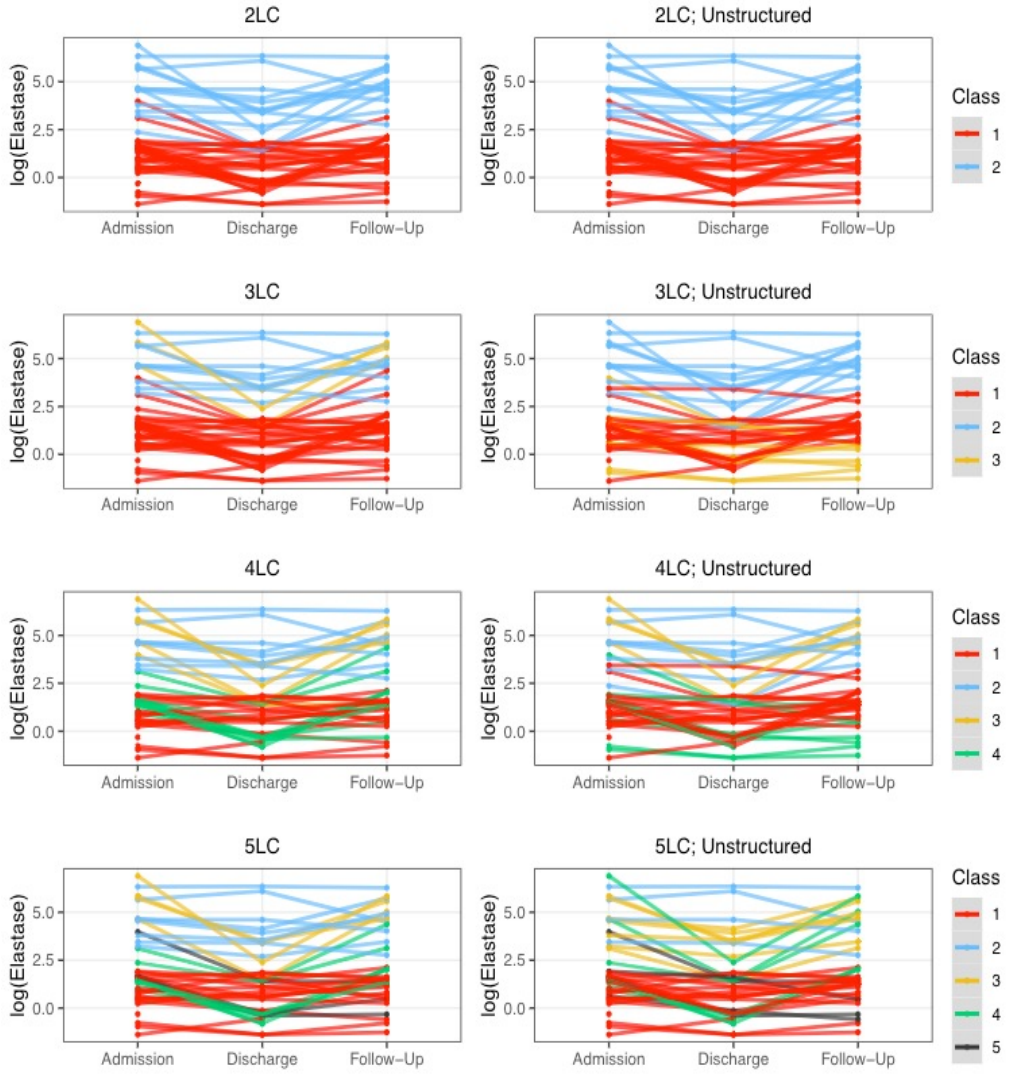
$$\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

LCMM Model Selection

- Model selection process
  - Considering a mixture of AIC, adjusted BIC, latent class balance, and visualizations to assess which model makes the most sense
  - Decision: 4 LC, diagonal variance-covariance structure for random effects

Table 2: Model comparison of latent class mixed-models fit with different numbers of latent classes

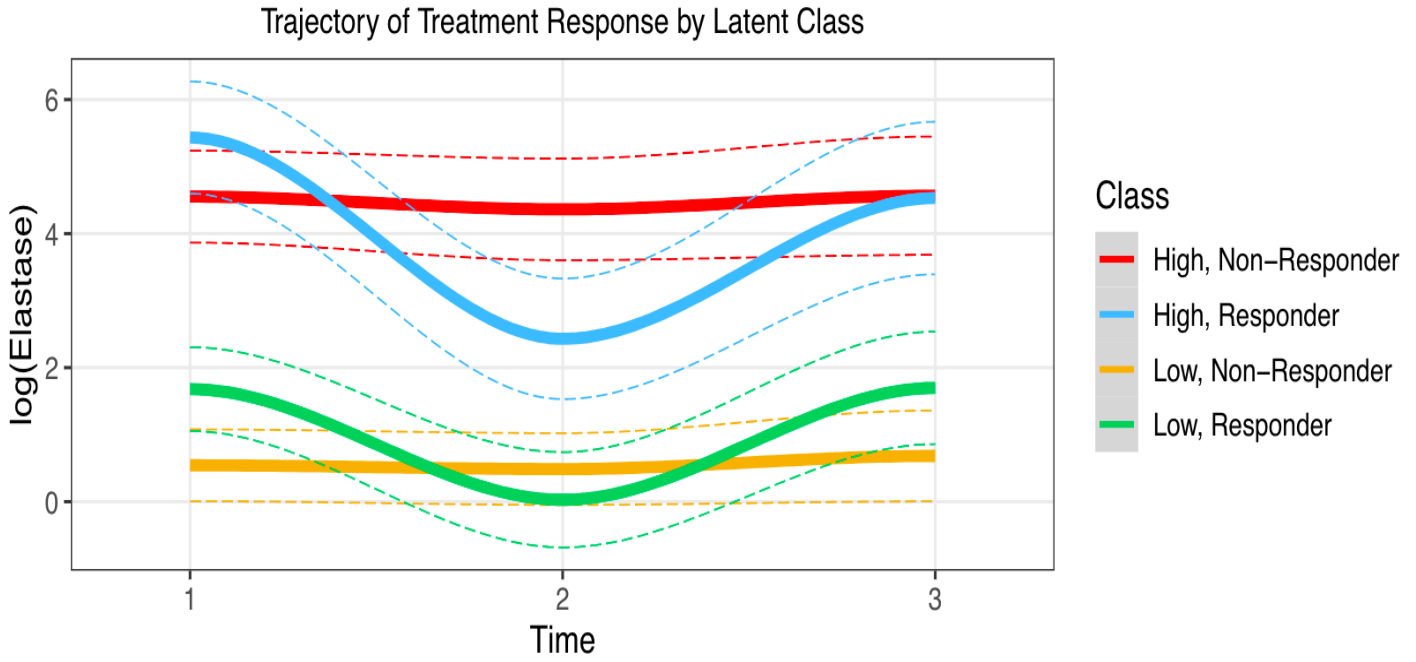
	Latent Classes	AIC	LC1	LC2	LC3	LC4	LC5
m2	2	390.5	66.7%	33.3%			
m2.unstr	2	394.4	66.7%	33.3%			
m3	3	391.7	69.2%	23.1%	7.7%		
m3.unstr	3	395.4	48.7%	30.8%	20.5%		
m4	4	392.3	41%	20.5%	12.8%	25.6%	
m4.unstr	4	396.4	48.7%	20.5%	10.3%	20.5%	
m5	5	392.8	43.6%	20.5%	10.3%	17.9%	7.7%
m5.unstr	5	404.4	43.6%	10.3%	17.9%	17.9%	10.3%



LCMM Selected Model

```
## Variance-covariance matrix of the random-effects:
##               intercept timeDischarge timeFollow-Up
## intercept       0.77158
## timeDischarge   0.00000       0.16506
## timeFollow-Up   0.00000       0.00000        0.6319
##
##                                coef      Se
## Residual standard error:  0.28291 0.14426
```

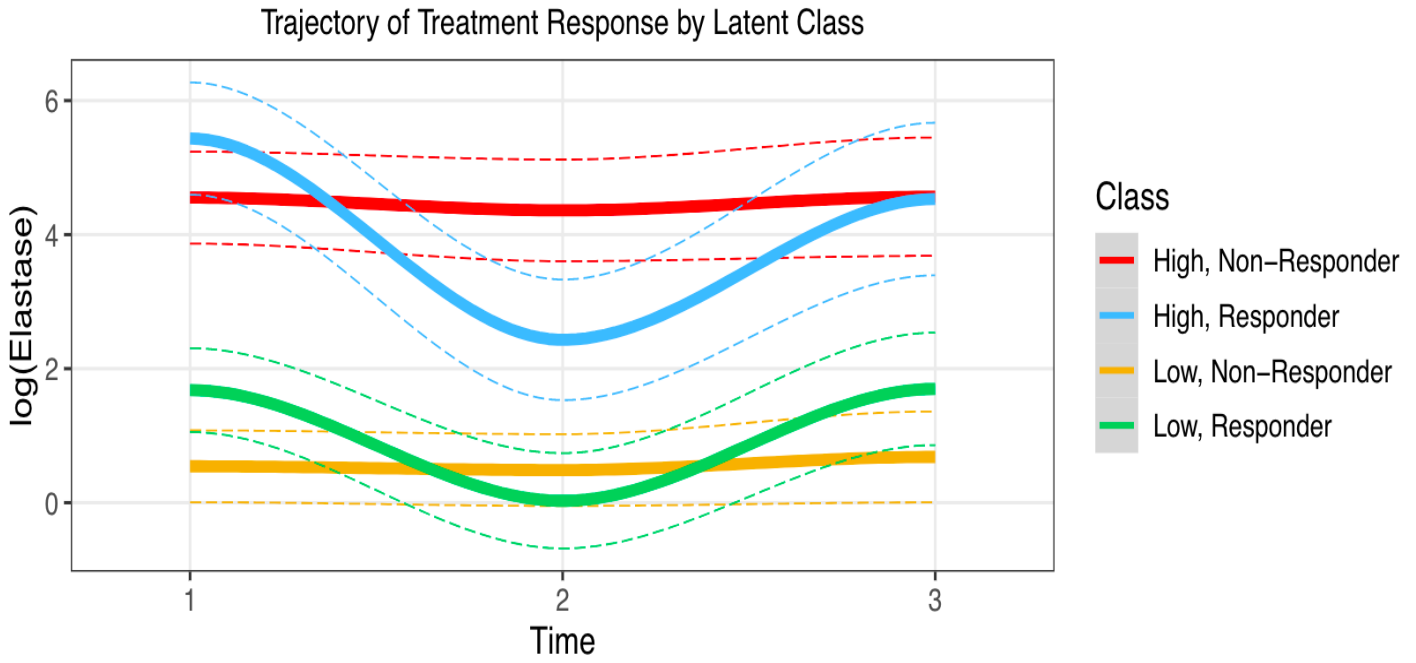
LC	Baseline Elastase	Treatment Response
1	High	Non-Responder
2	High	Responder
3	Low	Non-Responder
4	Low	Responder



- LCMM model also reports the probability of an observation belonging to each latent class

LCMM Selected Model Inference

	coef	Se	Wald	p-value
intercept class1	0.54333	0.27340	1.987	0.04689
intercept class2	4.55230	0.34978	13.015	0.00000
intercept class3	5.43403	0.42725	12.719	0.00000
intercept class4	1.68009	0.31923	5.263	0.00000
timeDischarge class1	-0.05622	0.17430	-0.323	0.74704
timeDischarge class2	-0.19015	0.20726	-0.917	0.35891
timeDischarge class3	-3.00442	0.25949	-11.578	0.00000
timeDischarge class4	-1.65182	0.24738	-6.677	0.00000
timeFollow-Up class1	0.13986	0.24623	0.568	0.57004
timeFollow-Up class2	0.01391	0.32201	0.043	0.96555
timeFollow-Up class3	-0.90283	0.41389	-2.181	0.02916
timeFollow-Up class4	0.01850	0.32668	0.057	0.95484

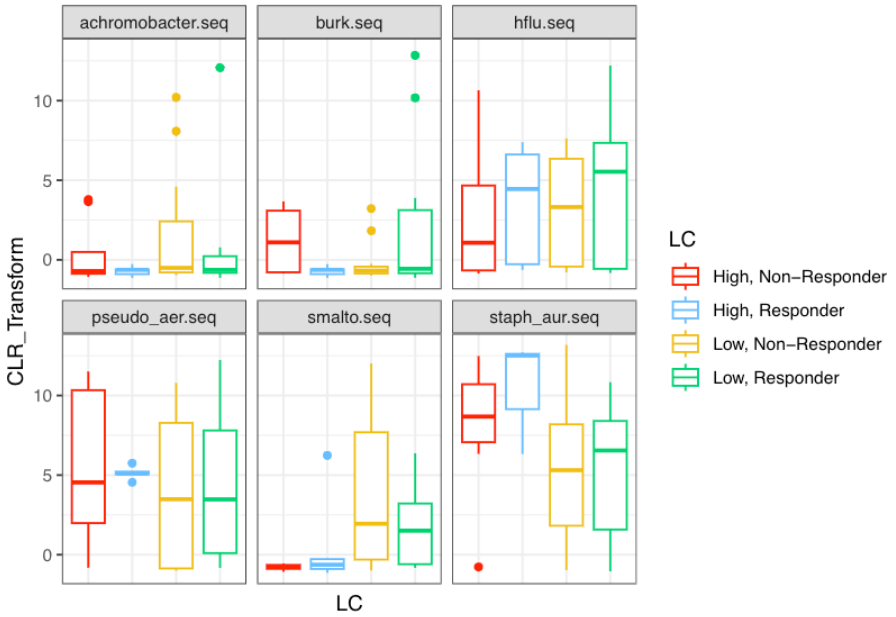


Post-Hoc Analysis

- Final Step
  - Use something like multinomial regression or binomial regression to assess baseline risk factor differences between latent classes; can use table 1 and data visualizations also to guide comparisons

Table 1: Demographics for study population by hard-drug use

	High, Non-Responder	High, Responder	Low, Non-Responder	Low, Responder
n	8	5	16	10
Female (%)	4 (50.0)	0 (0.0)	10 (62.5)	5 (50.0)
Age	15.50 [13.00, 20.00]	16.00 [12.00, 18.00]	15.73 [11.93, 19.90]	17.77 [12.91, 22.05]
BMI	19.98 [16.97, 23.00]	18.51 [17.95, 20.48]	19.57 [16.42, 28.70]	19.89 [16.61, 25.14]
Mutations				
0 F508del	2 (25.0)	0 (0.0)	1 (6.2)	1 (10.0)
1 F508del	4 (50.0)	2 (40.0)	5 (31.2)	1 (10.0)
2 F508del	2 (25.0)	3 (60.0)	10 (62.5)	8 (80.0)
Virus Present (%)	1 (12.5)	0 (0.0)	7 (46.7)	2 (28.6)
Pseudomonas	4.54 [-0.81, 11.51]	5.09 [4.54, 5.74]	3.48 [-0.99, 10.78]	3.48 [-0.82, 12.23]
Staph Aureus	8.68 [-0.77, 12.47]	12.51 [6.33, 12.70]	5.31 [-0.96, 13.17]	6.55 [-1.04, 10.83]
Achromobacter	-0.71 [-1.07, 3.78]	-0.63 [-1.14, -0.28]	-0.50 [-0.95, 10.20]	-0.63 [-1.12, 12.07]
HFlu	1.07 [-0.86, 10.65]	4.45 [-0.63, 7.38]	3.31 [-0.79, 7.63]	5.54 [-0.82, 12.20]
Smalto	-0.79 [-1.07, -0.56]	-0.63 [-1.14, 6.23]	1.94 [-0.99, 12.01]	1.51 [-0.82, 6.37]
Burkholderia	1.10 [-0.86, 3.67]	-0.63 [-1.14, -0.28]	-0.68 [-0.99, 3.22]	-0.56 [-1.12, 12.85]



Limitations

- In general, we should be careful about using latent class assignments as actual classes
- Sample size is a serious limitation here
- Results are exploratory; shouldn’t necessarily use the results for inference