* Mycobacterium avium complex (MAC)
* Disproportionately affects middle-aged and elderly, resistant to treatment
* Can develop in preexisting abnormal lungs (bronchiectasis or emphysema) or normal lung tissue
  + Bronchiectasis - a chronic lung condition where the walls of your airways (bronchi) widen and are thickened from inflammation and infection
  + Emphysema – lung disease that damages the alveoli in your lungs
* Fibrocavitary disease – severe form of NTM infection that occurs when the MAC bacteria creates a cavity or hole in the lung tissue (this is the “classic” NTM infection, described principally as upper lobe fibrocavitary disease with underlying emphysema)
* However in last three decades, nodular-bronchiectasis pattern (“non-classic infection”) has been increasingly observed. Occurs in any lobe, but more severe in right middle lobe (RML) and lingula, and possibly right upper lobe, manifesting as advanced bronchiectasis and atelectasis
* In studies, NTM lung disease appears to be more severe in RML and lingula
  + The reason for this is not known
* Acquisition:
  + Inhalation of NTM-contaminated water, biofilm, or soil aerosols
  + Microaspiration (from above – swallowing dysfunction, or below-GE reflux)
* Why more apparent on right?
  + In adults, right mainstem bronchus makes a less acute angle at its takeoff from the trachea, and is shorter than the left mainstem bronchus (2.5cm vs 5cm on average)
  + About equal in children, however
  + Secretions traveling down are more likely to go into these areas?
  + RML bronchus is more narrow, might be harder to clear airway with forced exhalation
  + RML and lingula are near the heart, the constant beating might cause microatelectasis (small passages are blocked or pressure is applied to outside of lung), might set up a nidus of infection (place in which bacteria have multiplied) that can spread
  + The “tapping” from the heart beat might skew macrophages in those areas to other types of macrophages (without other underlying causes, might not be super convincing, since these would be present in everyone
* People with primary ciliary dyskinesia (genetic mutations that affect the tiny hairline cilia in the lungs, nose and ears, impairing their ability to remove germs and pollutants, and allowing mucus buildup and infection) more predisposed to bronchiectasis and NTM lung disease

Notes from meeting

* Drug resistant TB
* NTM
* Look for disease in lungs
  + Large nodules, ground glass opacities
* To account for 6 measures per subject, add random intercepts
  + For thesis, try other types of correlation
  + Maybe ways to play with correlation structure
  + Can do comparison between lobes
* Scoring 0, 1, 2, 3
  + 0
  + 1: <25%
  + 2: 25-50%
  + 3: >50%
  + Take into account all radiologists (2 people scoring, 3 if tiebreak needed)
* Ordinal logistic regression: basic approach assumes proportional odds (this is a big assumption)
* Ordinal logistic regression
  + 3 intercepts
  + Treat CT findings as separate outcomes (separate model)
  + Repeated findings are the different lobes
* ICC: variance between raters
* 2 raters and ways to account for repeated measures: this will be the difference between preliminary data and my project

2025 – re-reading notes

* Matt’s work had 71 subjects, mine might have up to 164
* Model each lung pattern separately
* GLMM, random intercept
* Other ways to account for repeated measures
* Would make sense to do AR1 if (?) over time
* Spatial correlation?
* Easiest is a random intercept (correlation would be same
* Need to account for correlation in some way
* Right df
* Two part model (0 + continuous)
* Steps
  + Refit
  + Any other ways to account for correlation other than random intercept
  + 2-part model (0+continuous)
* Cumulative logit: odds interpretation
  + Odds of having higher score for particular lobe compared to other is…
* Proportional odds assumption (should maybe test for when it doesn’t hold either
  + Odds interpretations are valid across all levels of covariates
* When i fit model, go in and interpret each of the estimates
* ICC -> these models not fitting error term, not really able to do this
* If you use pseudolikelihood, we could? (another approach, he uses quadrature, not pseudolikelihood)

Before meeting on 1/14

* Repeated measures: 6 lobes scored (within subject repeated measures)
* Also the 2 (or 3) raters
* Random intercept?
* Consider partial proportional odds if proportional odds fails

From my meeting on 1/14

* Nothing formal to present for March, just getting together; PPT slides, can make simple paragraphs, where I’m at and what our plan is, can mention data analyzed for previous abstract
* Doing definitive analysis and more data
  + Multiple raters
  + Comparing SAS and R
  + Ordinal logistic regression
  + Playing with different correlation structures
  + Testing proportional odds assumption (test with a more complex model that doesn’t require this assumption)

Meeting on 2/21

* glmmTMB?
* Frame the meeting as “we are working through stuff” (adding random effects, correlation structures) as opposed to “we are struggling”
* Rater as a fixed effect: inference based on two raters (not ideal?)
* GEE: compound symmetric correlation (? Is this what I have for the exchangeable correlation structure?)
* How does testing for proportional odds affect the results?
* Another research question: How consistent are the raters?
  + Random intercept helps compute interclass correlation (within subject and between subject correlation)
  + Might not have to worry about different raters for future studies if they are the same
* Start with descriptive approach
* SAS: GLIMMIX, NLMIX
* 10 slide powerpoint
* Mention the data set is not 100% final, but still give some quantitative results

Meeting on 3/6

* Should also include plural effusion and lymphadenopathy in the binary category (but these don’t have multiple scores per lobe)
* Further NTM notes
  + Found in environment (“slimy stuff” in water)
  + Inhaled, causing lung disease
  + Lung notes in general: more stress in upper lobe when breathing in. decreased bloodflow to upper lobes, so he would think there are more cavities here
* Might want to try different convergence methods (method = quad)
* LLS: left lingular, not left lower
* The species on my last slide should be NTM species, not MAC species
* If ICC close to one => consistent
* Maybe also see if rater-by-rater (i.e. one rater’s scores) models give same as the other rater and/or the two combined
* Spatial correlation?
* Report the size of the cavities, given in mm (he said he tried to change all cm measurements to mm, if we see something that is 1mm, this is unlikely, probably 1cm

Call with Ed on 4/10

* Abstract questions
  + Remove the “pilot study” portion; “in this study…”
  + Remove the consensus statement, since we are incorporating both raters in the analysis
  + Maybe focus just on bronchiectasis, tree in bud (ignore the centrilobular nodules), and cavities (can treat both thick and thin as one for this abstract)
* General questions
  + Pleural effusion and lymphadenopathy: these are outside lung, in chest. So this is a presence or absence
  + For mosaicism: don’t count the ones with NA (just perform analysis on ones with 0 or 1; he was saying they are interested in the proportion of 1/(0+1)
  + Pleural effusion is fluid around lung: L=left, R=right, B=both; this many L, R, B; yes/no analysis
  + More women than men: true that it affects more women than men in population (maybe 6 or 8 women to 1 man), but the discrepancy we see is probably due to referral bias (idk if that’s a real bias, I just said it, he didn’t say it), as most patients are referred to the study by their physicians
  + Rating: Ed confirmed Jiwan rated separately and then Valerie Hale rated again and he entered the data while she was doing so. No correspondence between the raters, they have probably never met each other

Meeting with Matt on 4/11

* Matt says he might want a draft of the written component (analysis mostly) a few weeks before our next committee meeting on 6/11
* For the “lit review” in the intro, I can touch on the review of the methods and application
  + 25-30 references sounds sufficient to him, and less might not really be a big deal
* For the GGO model that we are having trouble fitting, try to fit a logistic regression model instead of the ordinal
* In general, go with what works and we are able to fit for modeling
* For ones that do work, do a comparison between the ones that fit random rater and ones that don’t
  + Maybe compare the lobe comparisons (rater is not main thing of interest
  + For ones that don’t work, see if taking out fixed rater will help
  + For ones that won’t work with the fully fledged model, go with what works, as this probably won’t affect the more general inference we are doing
  + By taking it out of the random term, how does it affect lobe comparisons?
* Multiple testing correction for the pairwise comparisons
  + FDR (false discovery rate)
  + Benjamini Hochberg
  + Or maybe Tukey-Kramer
  + These are controlling false discovery rate (not alpha)
  + Check to see if proc glimmix or lsmeans has this (adjust = …)
  + FDR will be more liberal, i.e. more significances
* For the cavitary ones, use Fisher’s (exact?) test for abstract
  + For thesis, we should be using the model. Probability of cavity is more meaningful than the count of cavity (?)
* As I’m running models, send him summary of models and results (keep it concise)