

# **Fail Fast: Early Prediction of Outcomes and Optimizing Medications in IBD**

Peter D.R. Higgins, MD, PhD, MSc (CRDSA)

Director, IBD Program

Director, Clinical Trials Support Unit

University of Michigan

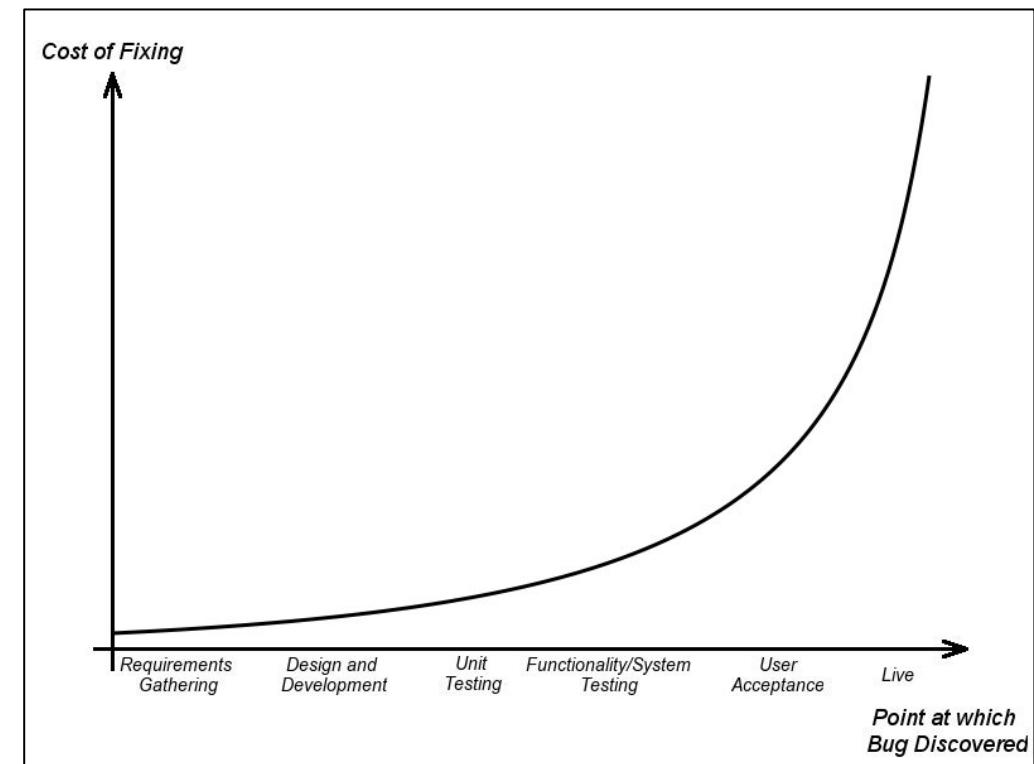


# Disclosures

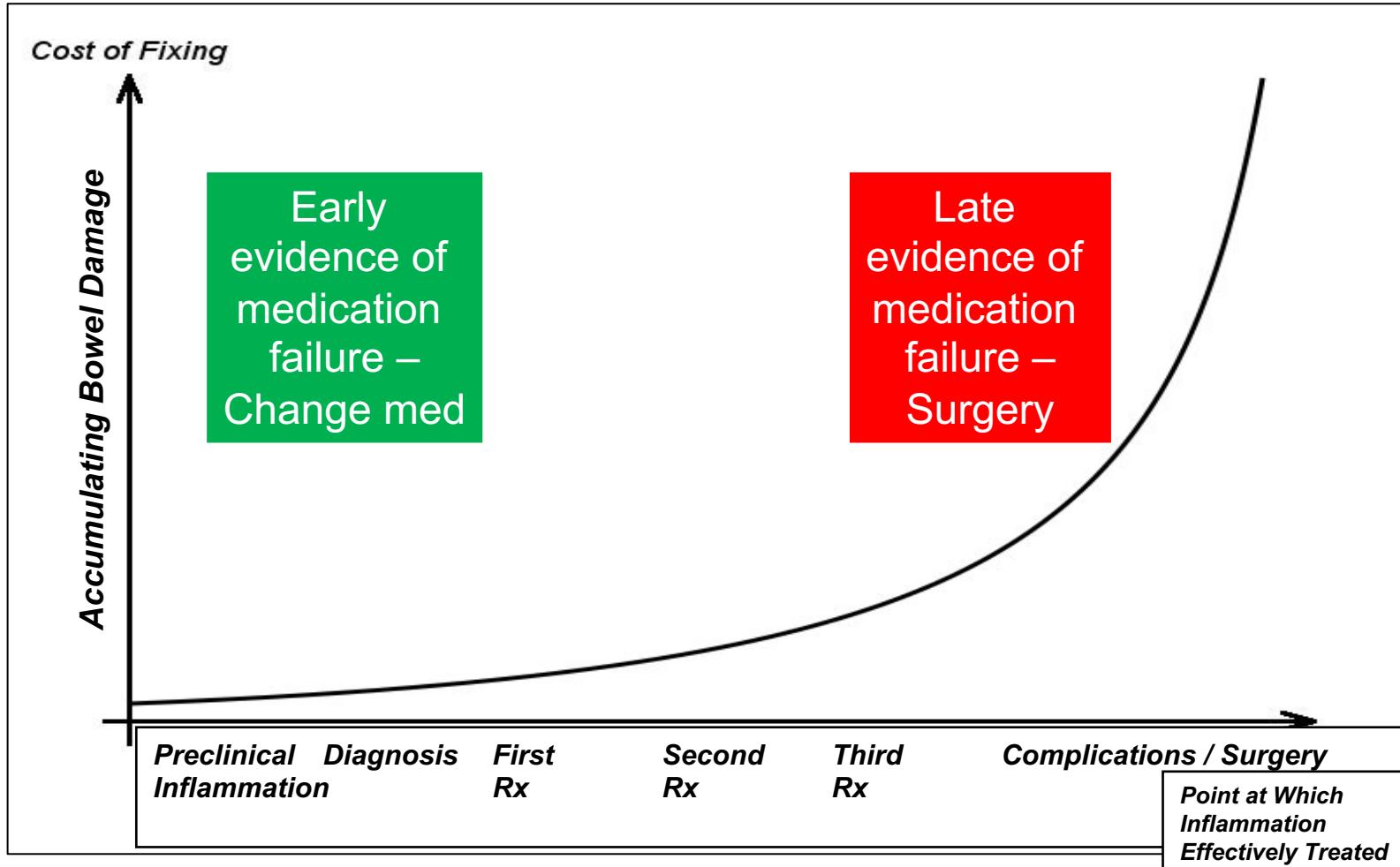
- Research Funding:
  - NIH
  - CCF
  - Broad Foundation
  - Abbvie
  - Ascentage Pharma
  - Siemens
- Consulting
  - Lycera, UCB, Eli Lilly, BMS
  - Amgen, PRIME Medical Education
- Speakers Bureau
  - GI Health Foundation
- Clinical Trial Site PI:
  - Abbvie
  - Amgen
  - AstraZeneca
  - BCBS
  - Celgene
  - Genentech
  - Gilead
  - Janssen
  - MedImmune
  - Nestle
  - Pfizer
  - Salix
  - Seres
  - Takeda
  - UCB

# What is “Fail Fast”?

- A concept from agile software development
- It is better for your code to fail early and explicitly than to quietly continue to process and crash later
- It is easier and cheaper to fix a bug early than it is to fix it later
- This works surprisingly well as an analogy for IBD care



# Failing Fast in IBD



Steroids and placebo effects can conceal early evidence of medication failure

# Speed of Rx Failure in IBD

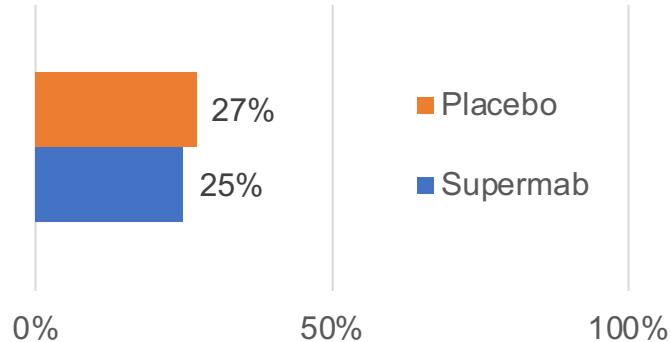
- Clinicians can be pretty slow to recognize therapeutic failure
- Thiopurines, methotrexate – wait 12-16 weeks
  - Dose adjust?
- Vedolizumab – wait up to 40 weeks?
  - Go to q 4 weeks?
- Ustekinumab
  - How long? Go to q 4 weeks?
- Can spend a lot of time rearranging deck chairs
- Long-term consequences from resulting **bowel damage**



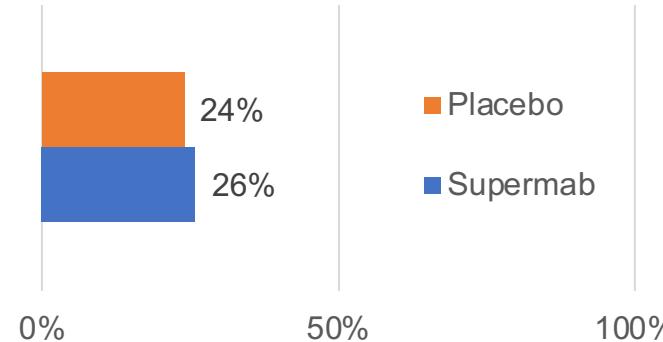
# Clinical trials with **SUPERMAB**

- Imagine a new biologic for IBD named Supermab
- 2 RCTs are conducted

Trial 1: Clinical & Endoscopic Remission



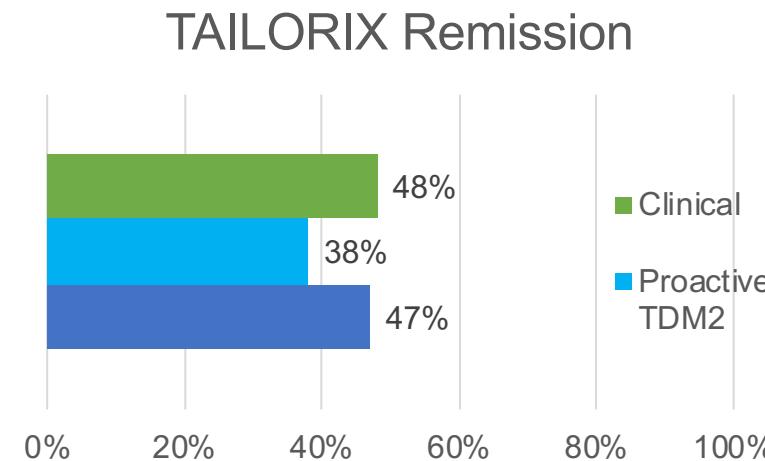
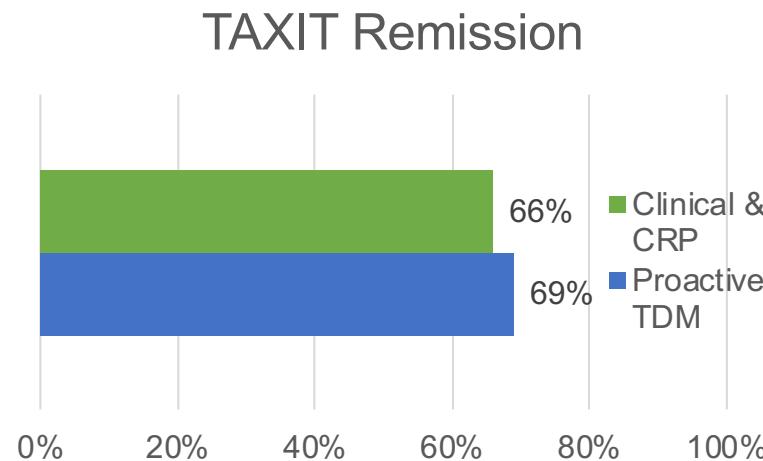
Trial 2: Clinical & Endoscopic Remission



Would you prescribe this drug for IBD?  
Should the FDA approve this drug for IBD?

# Notes on Proactive TDM in IBD

- 3 prospective RCTs of thiopurine metabolite monitoring and adjustment have failed to produce better outcomes
- 2 prospective RCTs (TAXIT, TAILORIX) of anti-TNF trough level monitoring and adjustment have failed to produce better outcomes



AGA Guideline:  
“Prospective studies  
have failed to show  
clinical benefits of  
proactive TDM”

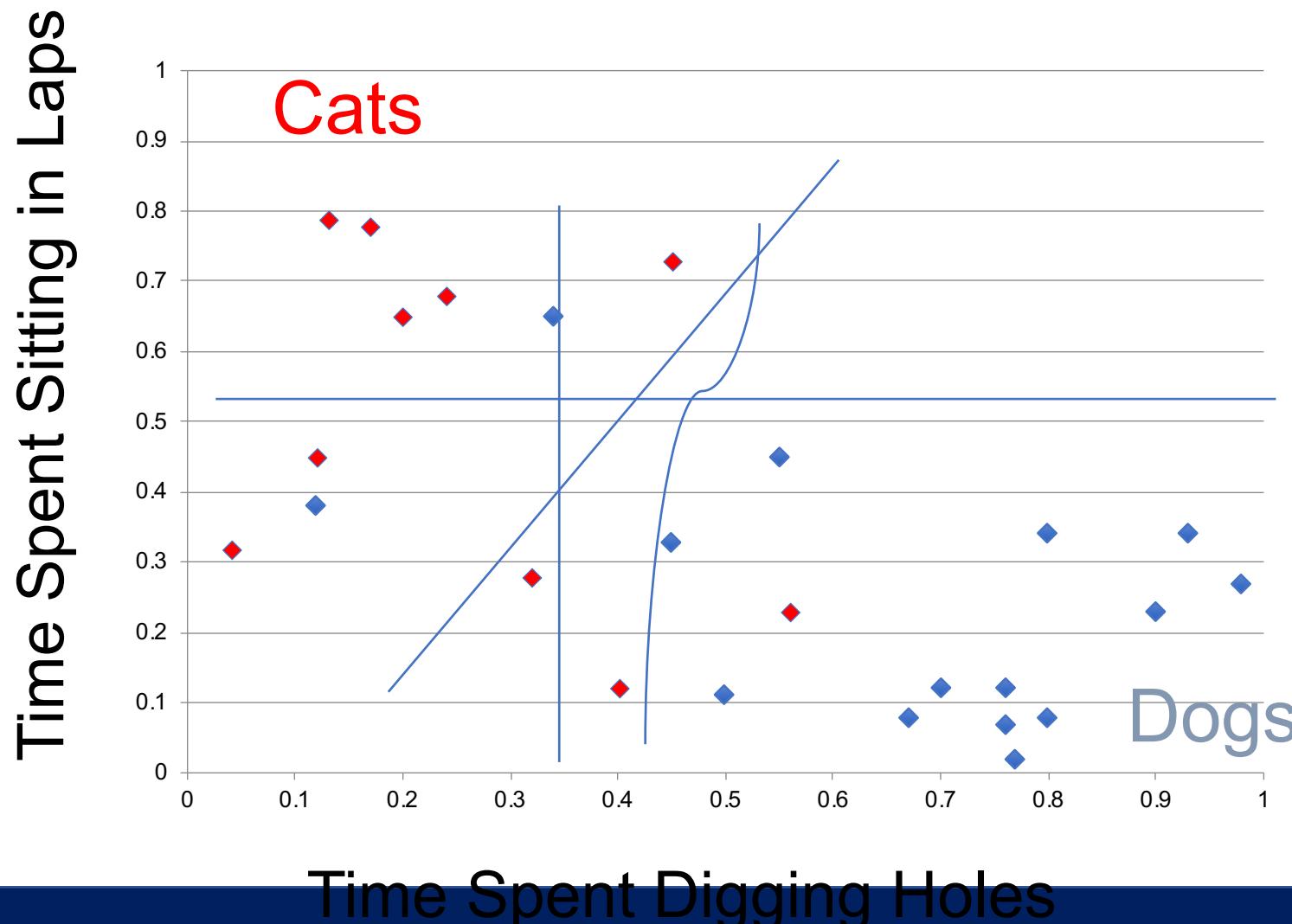
# Can We Fail Faster?

- Can we predict early failure or success with IBD medications?
  - Thiopurines
  - Vedolizumab in UC
  - Vedolizumab in CD
  - Ustekinumab in CD

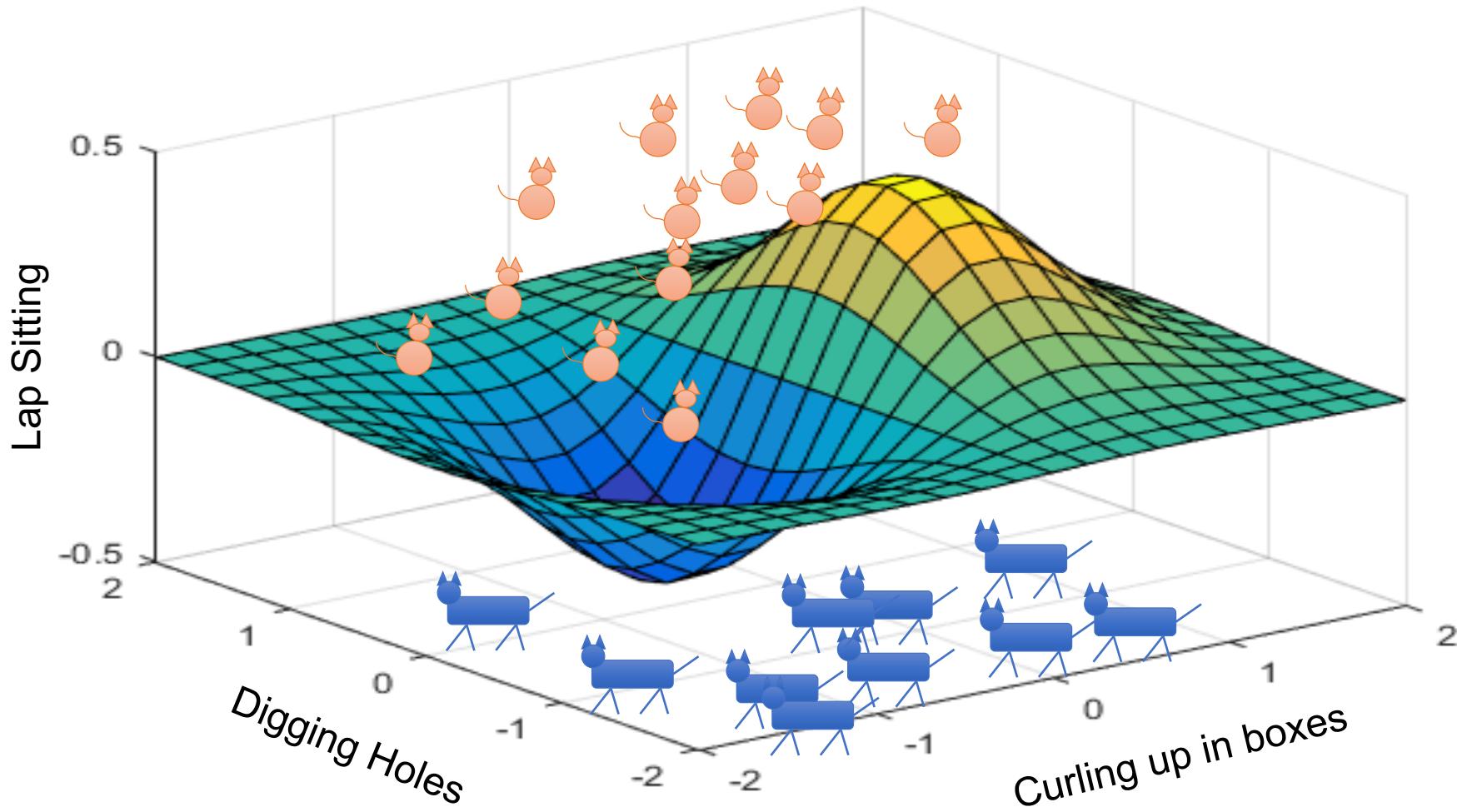
# Methodology

- Collect data on lab values associated with response/failure
  - High-quality, CLIA-certified data
- Collect biologic response outcomes, rigorously documented
  - Use clinical trial data when available – YODA, CSDR, etc.
- Split data into two random subsets
  - 70% for training a model
  - 30% for testing whether the model is generalizable
- Model-building
  - Multiple machine learning approaches, mostly random forest
  - If possible, identify a simple predictor

# Identifying Nearly Pure Splits



# 3 Variables = 3 Dimensions





# Thiopurine Machine Learning

Will this Patient benefit from this particular Therapy?

# The Problem

- Thiopurines have a poor therapeutic index
- Metabolism and efficacy are variable
- Metabolites are poor predictors of efficacy
- Expert physicians claim there are recognizable patterns in labs when patients are doing well on thiopurines
- **Can we optimize thiopurines with machine learning patterns?**

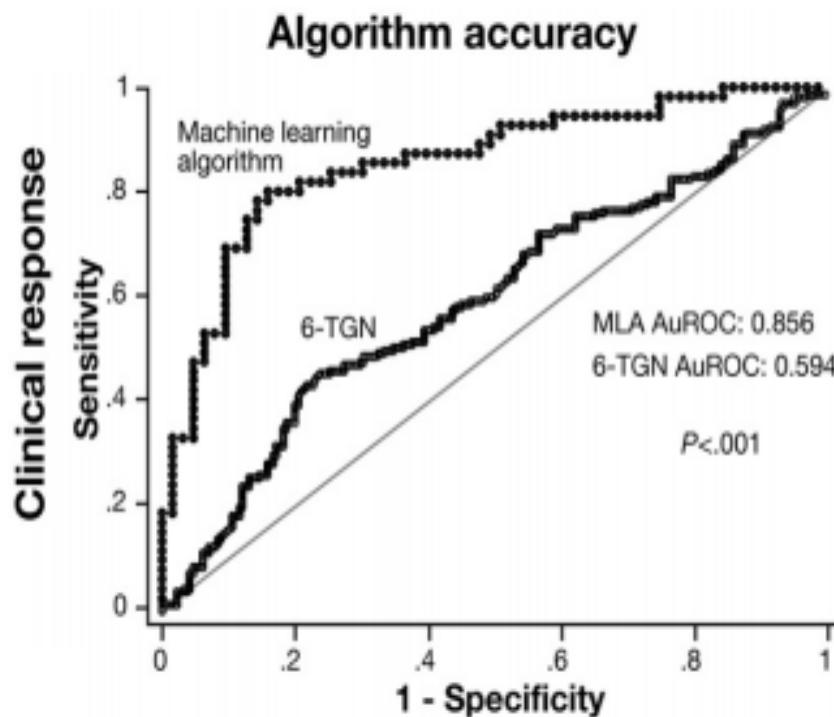
# Can Laboratory Patterns Distinguish Biologic Remission vs. no Remission?

## Results

# Background

- Machine learning algorithms are superior to 6-TGN in predicting clinical remission

Waljee, et al. Clin Gastro Hep.  
2010 8:143-50.



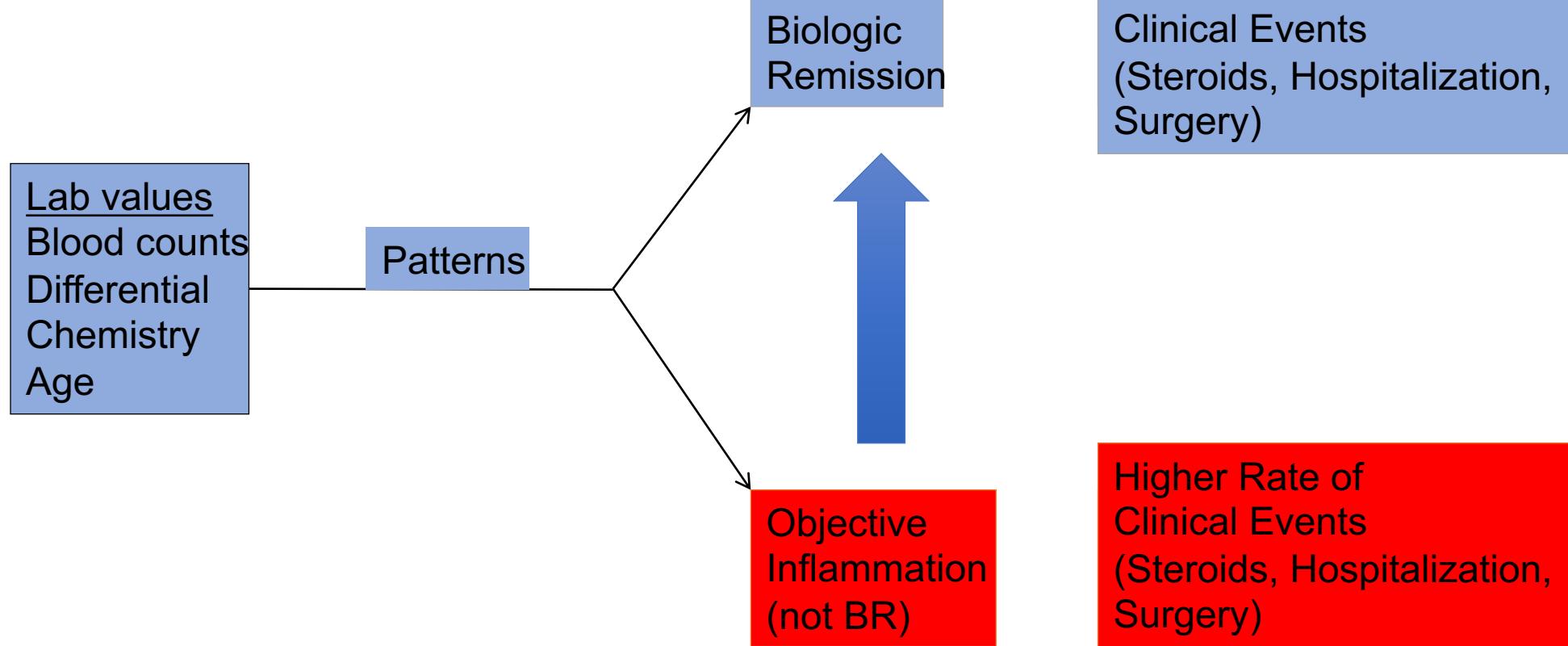
No data on biologic remission.

No data on clinical outcomes.

6-TGN optimization has failed to improve outcomes in 3 RCTs

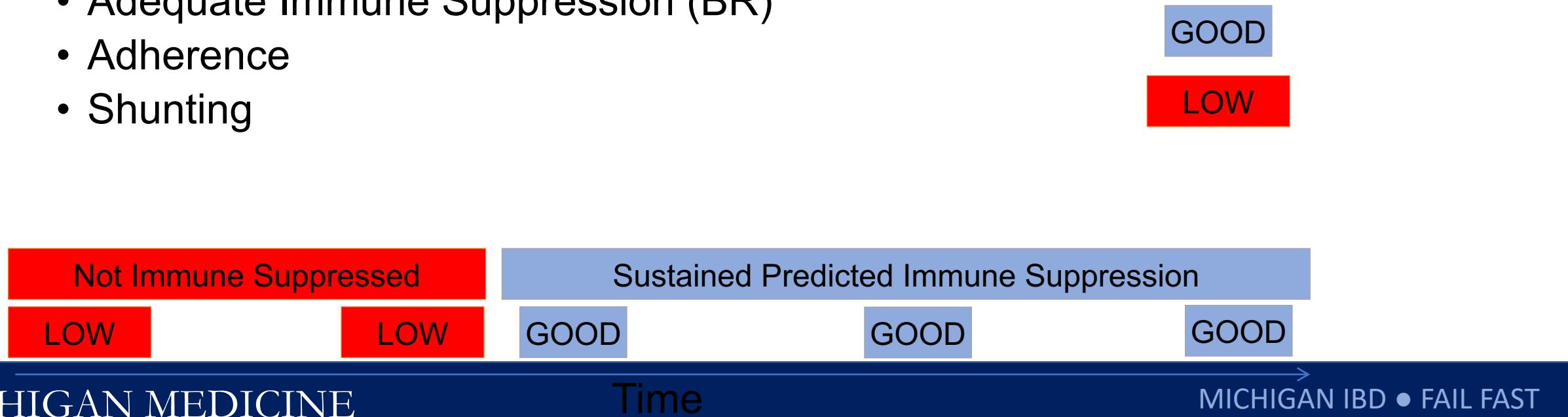
Reinshagen, et al. Clin Chem 2007 53:1306  
Gonzalez-Lama, Y., et al, APT 2011; 34: 544-54  
Dassopoulos, et al. APT 2014.39:163-175.

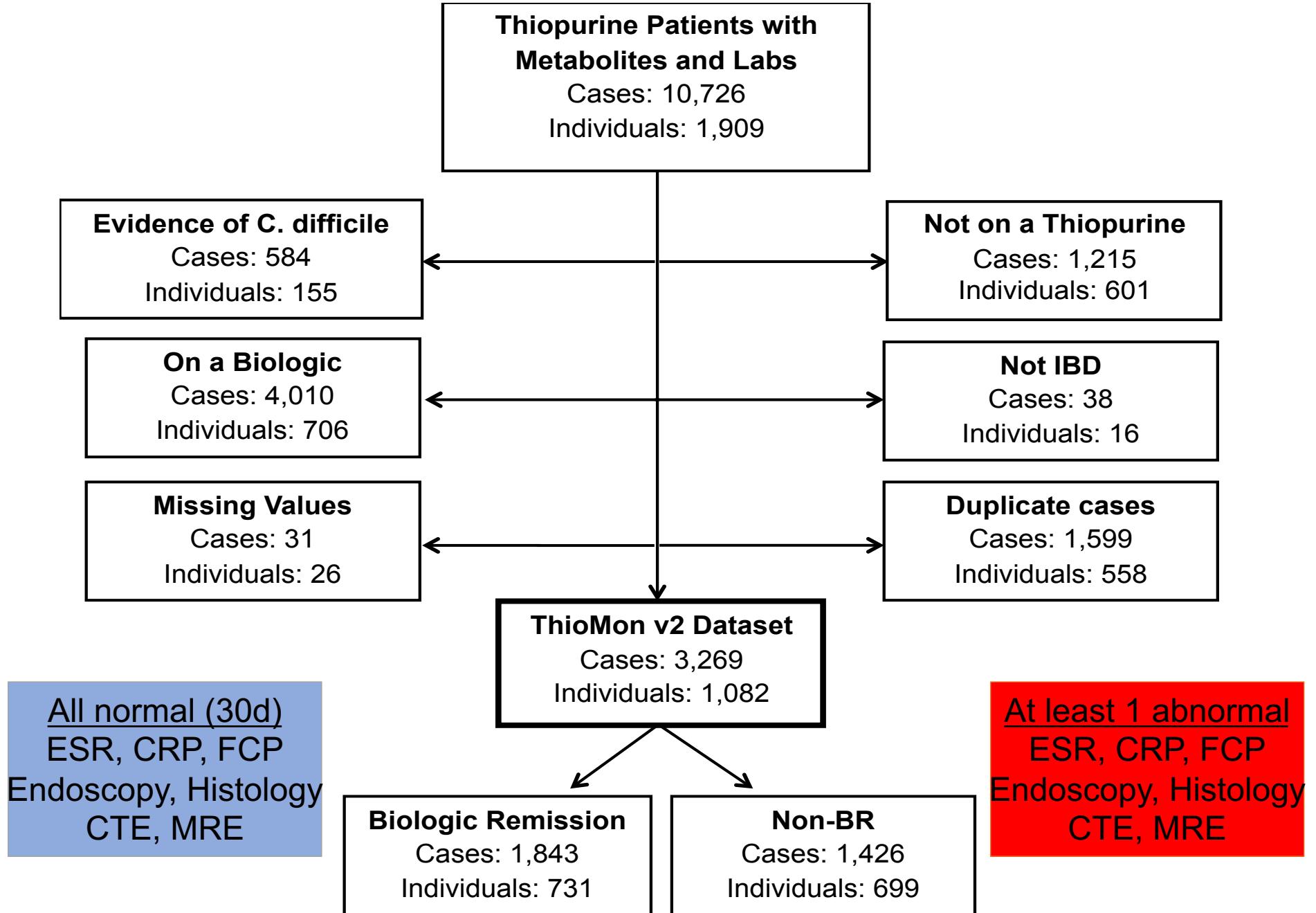
# 3 Hypotheses



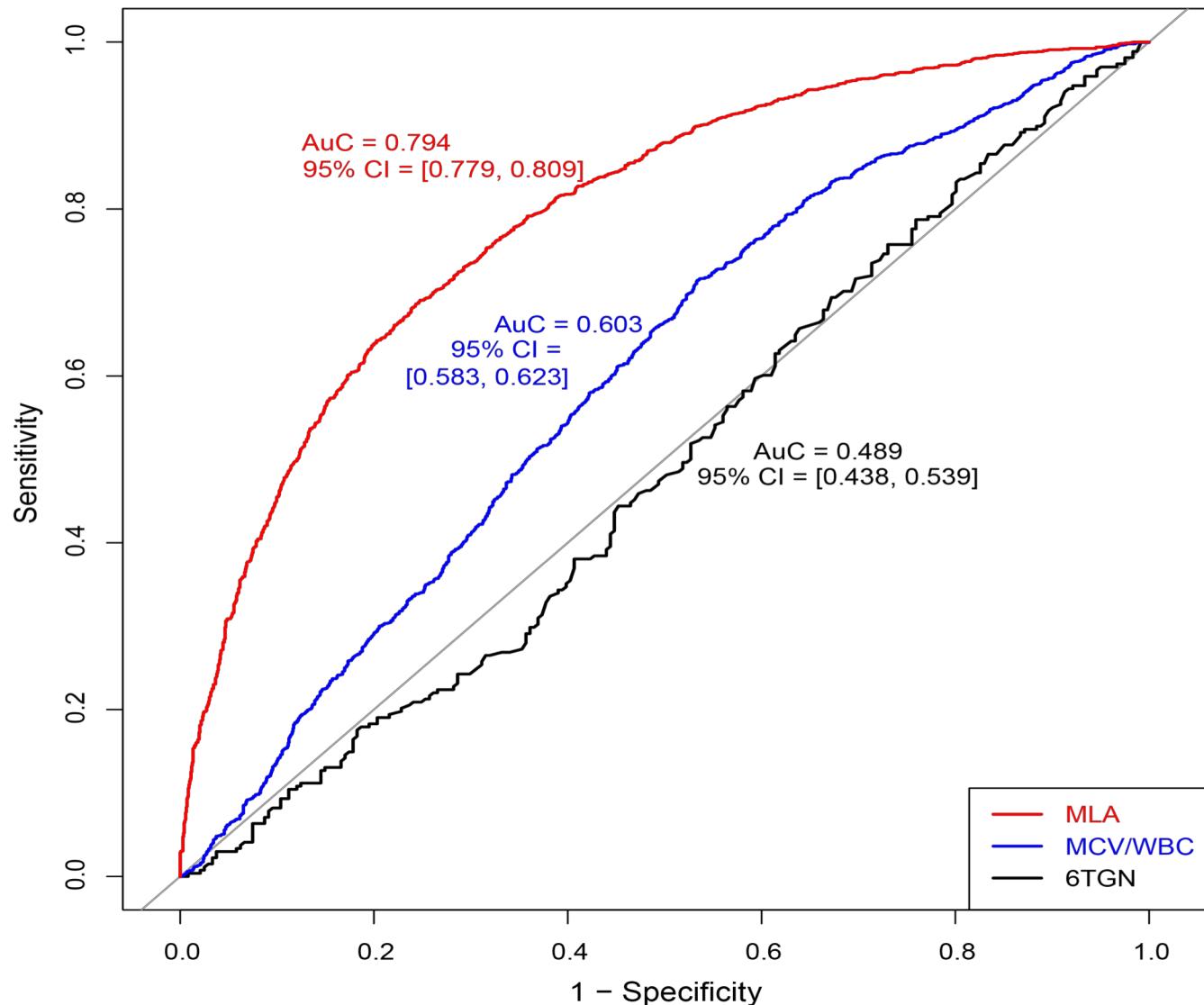
# Methods

- IBD patients on stable thiopurine dose  $\geq 4$  weeks
- Random Forest machine learning
  - Cross-validation on out of bag (OOB) sample
- Algorithms were developed for
  - Adequate Immune Suppression (BR)
  - Adherence
  - Shunting





## ROC Curves for Biologic Remission

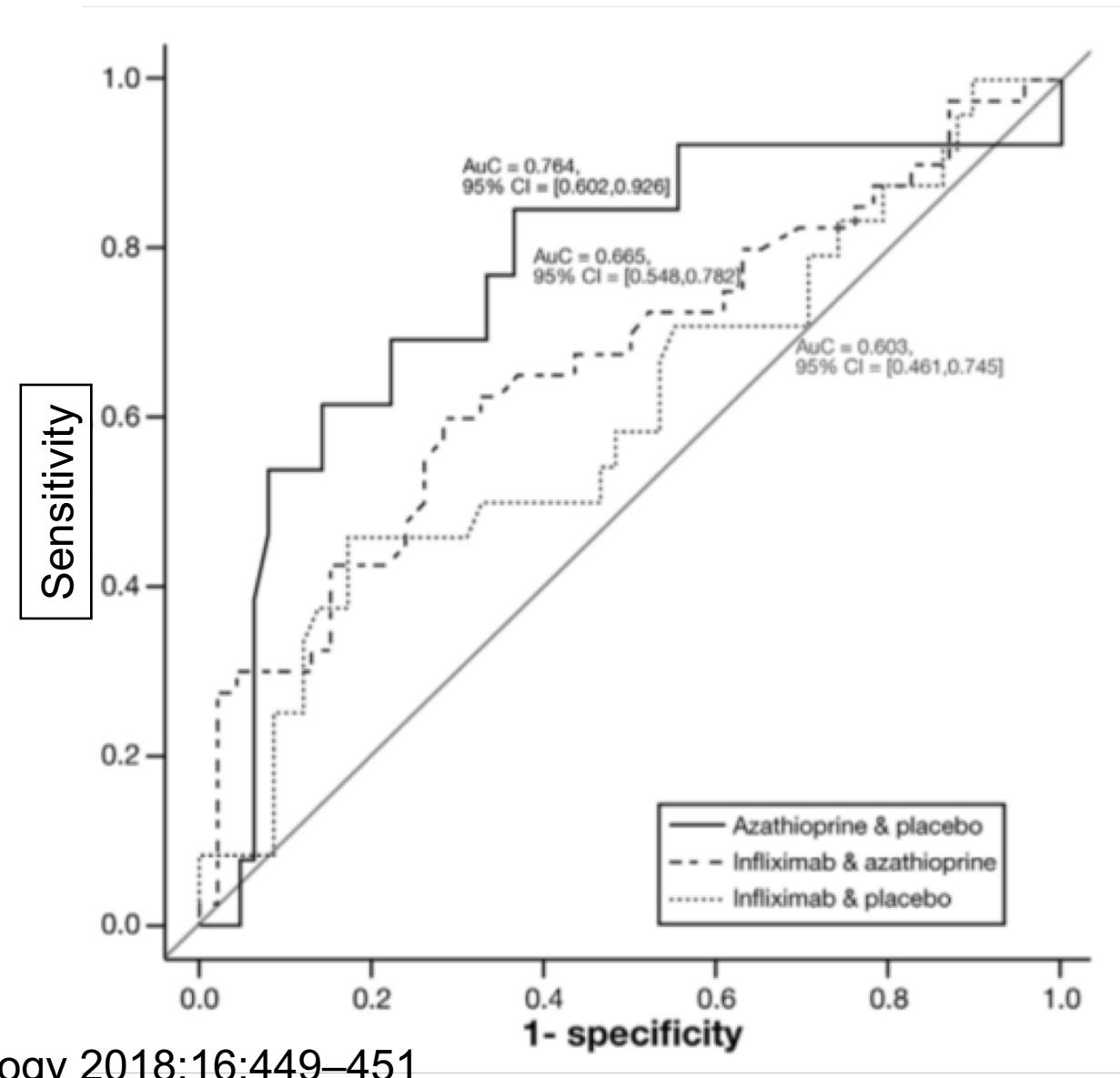


MCV/WBC ratio  
for everyday use

<u>values</u>	<u>ratio</u>
MCV 100	25
WBC 4	
MCV 80	20
WBC 4	
MCV 90	15
WBC 6	
MCV 75	13
WBC 7.5	
MCV 72	10
WBC 9	

# Does This Externally Validate?

- SONIC data
- Compared
  - Azathioprine
  - Infliximab
  - Aza + IFX
  - In Crohn's disease

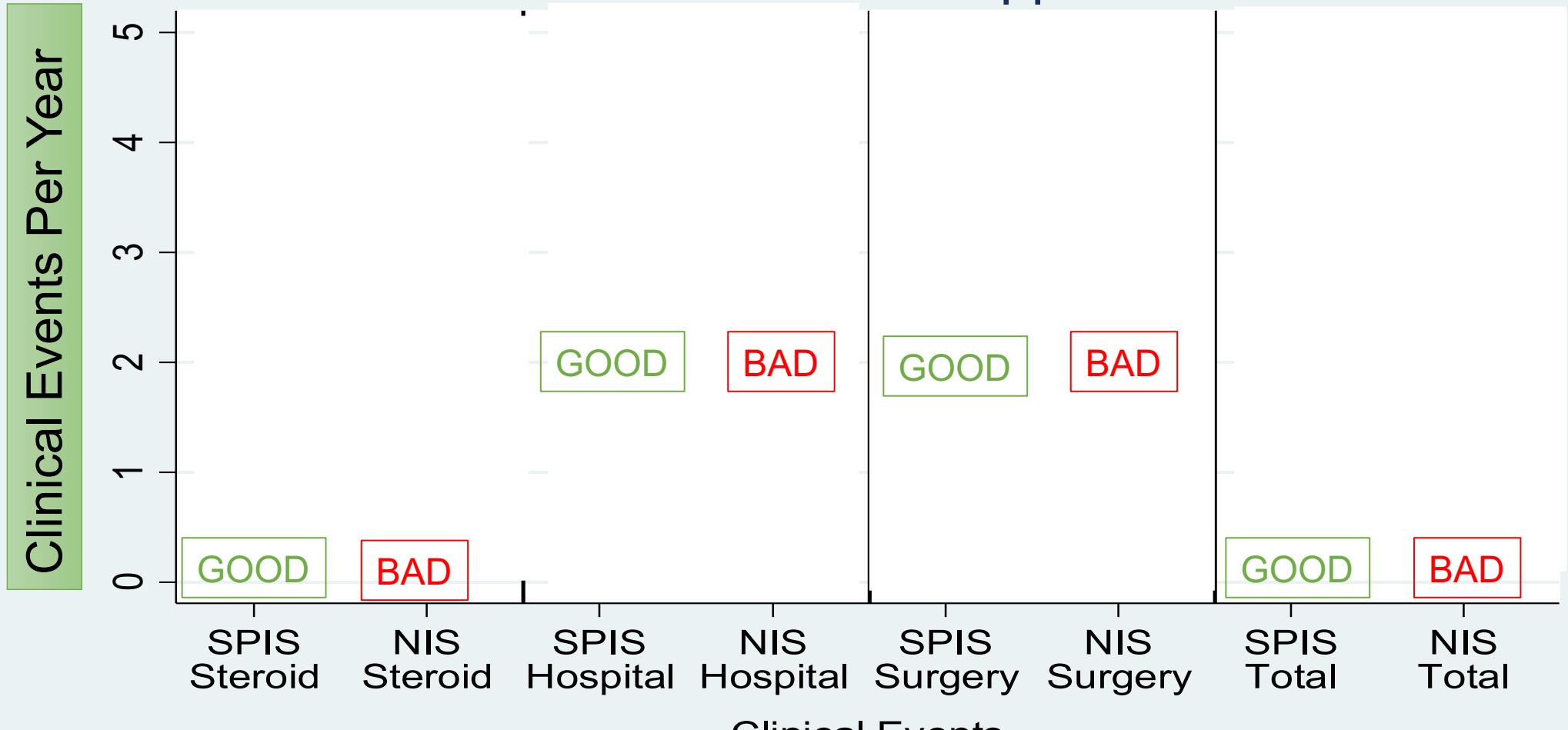


Waljee, et al. Clinical Gastroenterology and Hepatology 2018;16:449–451

## Results

Are Clinical Outcomes Better with  
Sustained Immune Suppression  
Pattern vs.  
Not Immune Suppressed?

# Clinical Event Rate with Sustained Predicted Immune Suppression N=274 vs. Non-Immune Suppression N=235



# RESULTS

If patients change from  
Not Immune Suppressed (BAD) to  
Sustained Immune Suppressed (GOOD),  
do Clinical outcomes improve?

# Effect of Achieving Sustained Predicted Immune Suppression with Thiopurines N=32 cases

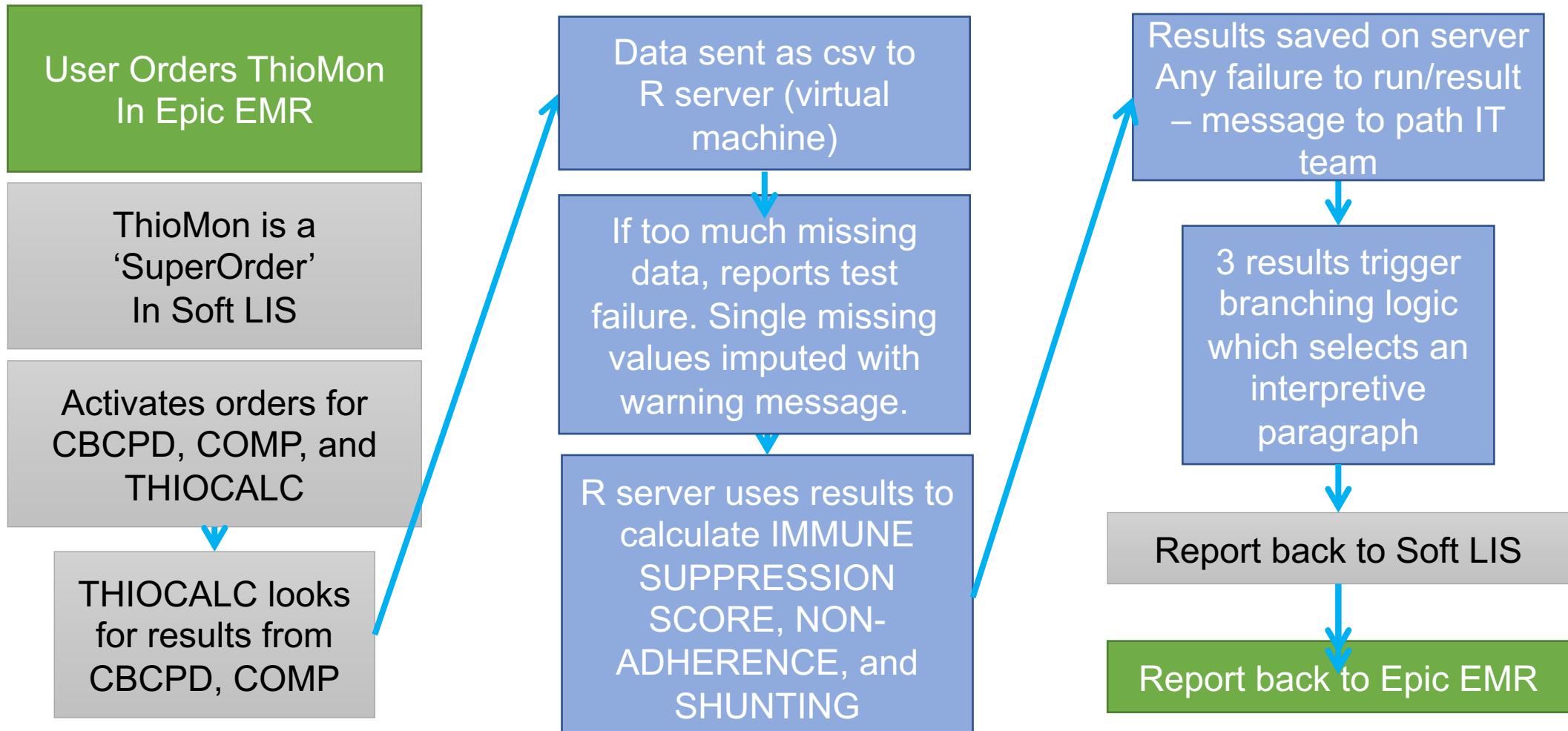


# Limitations

- These results are only as generalizable as the data from the phase 3 clinical trial that led to FDA approval.
- The simple MCV/WBC model was developed *post hoc*, and tested on the full data set, and would benefit from external validation.



# ThioMon implementation



# Result Reporting

- Provide limited interpretation with branching logic
  - Some providers will use this test < 10 times per year
  - Education from a year ago will **not** stick.

## Thiopurine Monitoring Test

Immunosuppression Score 102.3 (>100)

GOOD Result – patient has achieved effective immune suppression with thiopurines. If continued symptoms, consider infection, IBS, or drug failure necessitating a different class of therapy.

## Thiopurine Monitoring Test

Immunosuppression Score 92.7 (>100)

Shunting Score 91.1 (>100)

Non-adherence Score 94.4 (>100)

LOW Result – patient has not achieved effective immune suppression with thiopurines. No evidence of shunting or non-adherence. Consider increasing dose, adding allopurinol, or a different class of therapy.

# The Competition

- Metabolite testing from Prometheus Labs/Nestle
  - 6-TGN and 6-MMP are active and toxic metabolites
  - Measurable with HPLC, there is a CPT code
  - Mostly covered by insurance
  - NOT a good test – 3 prospective RCTs failed
  - But marketed very well
- These are the same people who can sell billions of chocolate bars contaminated with stale rice.



# Pathology dollars saved



- Previously ordered over 600 metabolite tests per year @ \$200 each
- Saved > \$120,000 per year in external costs
- Internal algorithm nearly free (virtual machine)
- Happy pathologists and accountants
  - A research project that actually **saved** money!



Dr. Jeffrey Myers  
Vice Chair of Clinical Affairs and Quality



# Vedolizumab Machine Learning

Will this Patient benefit from this anti-integrin Therapy?

# Vedolizumab in Ulcerative Colitis

- Vedolizumab is an effective, FDA-approved therapy for UC.
- Vedolizumab can be slow to produce complete remission
  - Only 16.9% in clinical remission at week 6.
- Vedolizumab, like other biologic medications, is very expensive.

List prices for a single infusion of Vedolizumab:  
- Illinois: \$9,973  
- New York: \$18,148

Date:	03/22/2017		
Payor:	PCP, DRG		
MRN:			
Expected	Payment	Adjustment	Balance
\$9,265.20	\$5,295.05	\$3,926.72	\$0.00
\$43.43	\$43.43		\$0.00
\$9,612.17	\$5,180.40	\$4,431.77	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,612.17	\$5,376.44	\$4,235.73	\$0.00
\$9,973.02	\$5,580.32	\$4,392.70	\$0.00
\$9,973.02	\$5,505.08	\$4,392.70	\$0.00
\$75.24	\$75.24		\$0.00
\$9,973.02	\$5,505.08	\$4,392.70	\$0.00
\$75.24	\$75.24		\$0.00
\$9,973.02			
\$9,973.02			

Feagan, BG, et al., NEJM 2013; 369: 699-7

# Decisions

- It can be difficult to decide what to do with the patient with limited early response

Keep dosing?



83.1 %

Shorten Interval?

Add a booster Rx?

Change to another Rx?

Aim: To predict at an early time point whether a UC patient on vedolizumab will achieve corticosteroid-free endoscopic remission at week 52.

# Methods

- For entry: active ulcerative colitis with Mayo 6-12 and sigmoidoscopy score of 2-3.
- Strict endpoint:
  - Week 52 Corticosteroid-free endoscopic remission:
    - No use of systemic steroids nor budesonide at week 52.
    - Mayo endoscopy score 0-1 at week 52

Feagan, BG, et al. NEJM 2013; 369: 699-710.

# Methods

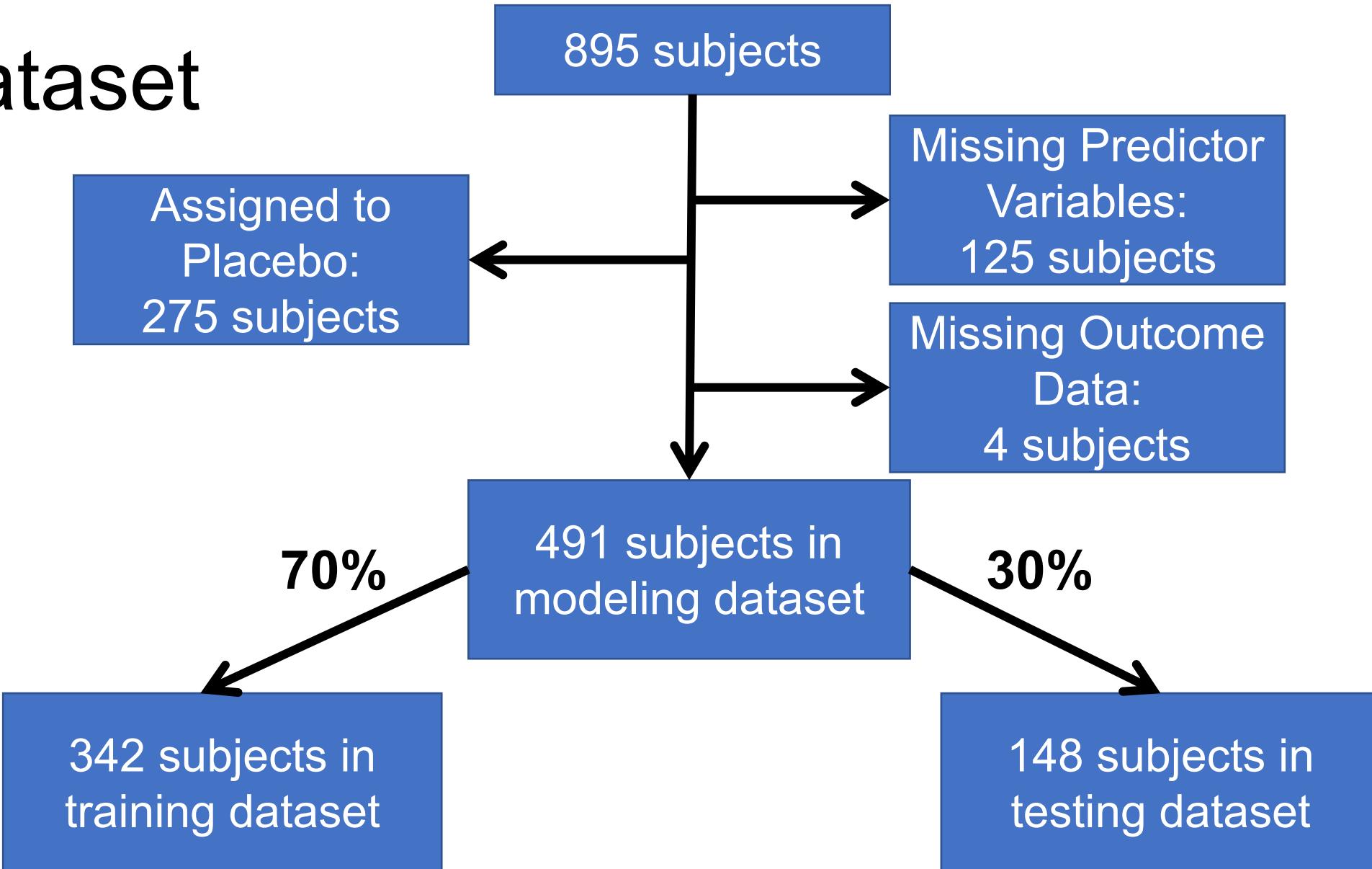
- Exclusions: subjects randomized to placebo, OR with missing predictors or outcome data
- Randomly divided complete data into 70% model training set and a 30% model testing set
- Trained model with 1000 RandomForest trees on data from:
  - Through week 6 (before 3<sup>rd</sup> dose of Vedolizumab)

# Predictor Variables Tested



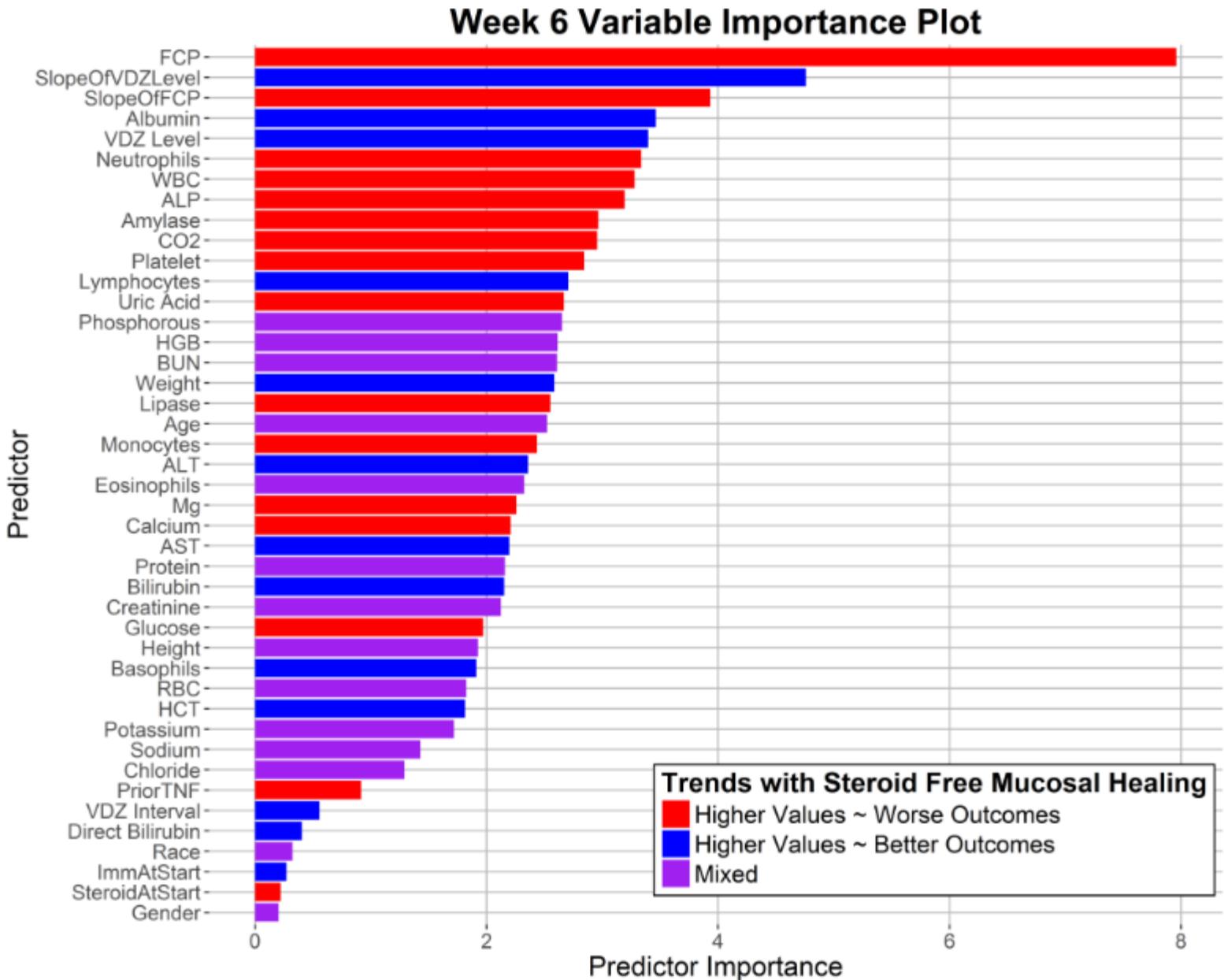
- All components of complete blood count with differential (0,6)
- All components of comprehensive chemistry panel (0,6)
- Uric acid, fecal calprotectin (0,6)
- Vedolizumab level at weeks (2, 6)
- Age, gender, race, height, weight
- Immunomodulator at entry, steroids at entry, prior anti-TNF
- Assigned vedolizumab interval (4 vs. 8 weeks)
- *Disease extent, baseline sigmoid activity, disease duration*
- Slope, mean, min, max, acceleration of all lab values

# Dataset



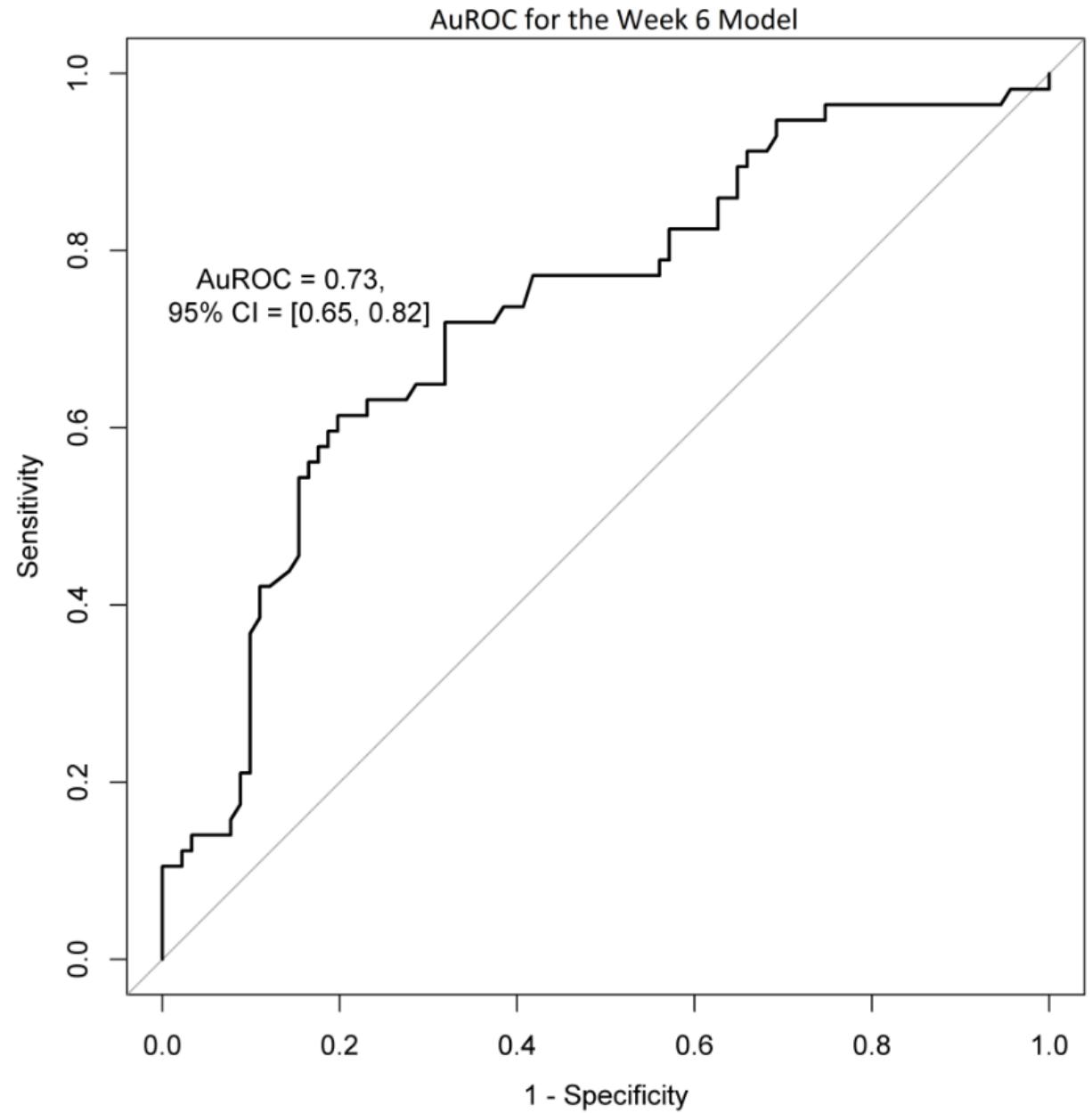
# Results

- Variable Importance at Week 6

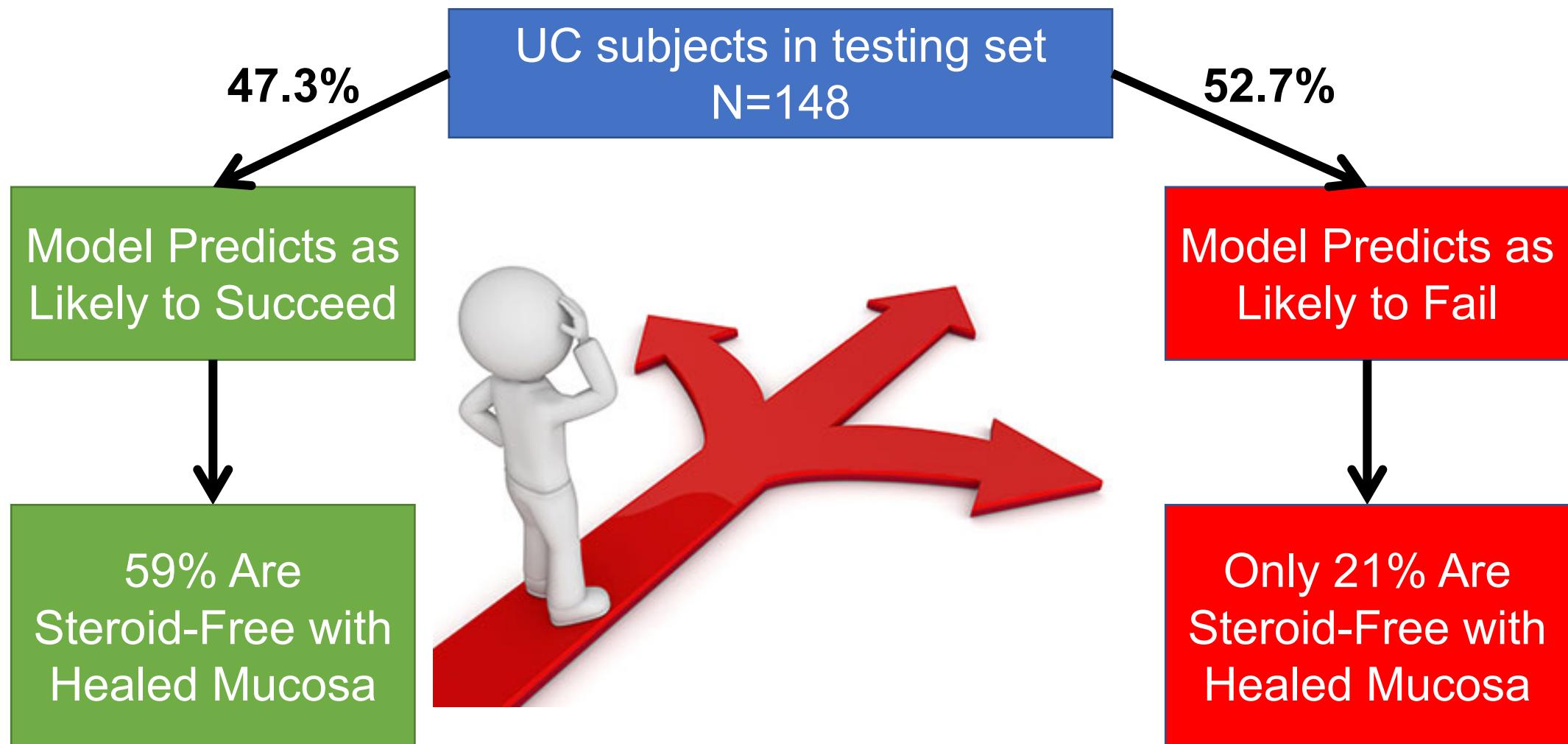


# Results

- Week 6 Model
- Accuracy on Testing Set



# Prediction Success with Week 6 Model



# Challenges to Applying these Results

- Need labs at multiple time points – complicated logistics
- Complex IT Infrastructure required to
  - Maintain privacy when sending health data
  - Link labs from multiple time points.
  - Run random forest models
- Is there a simpler way?
- Propose a simpler model *post hoc*
  - Single model (avoid multiplicity of testing)
- Test on full data set

Week 6 Fecal Calprotectin

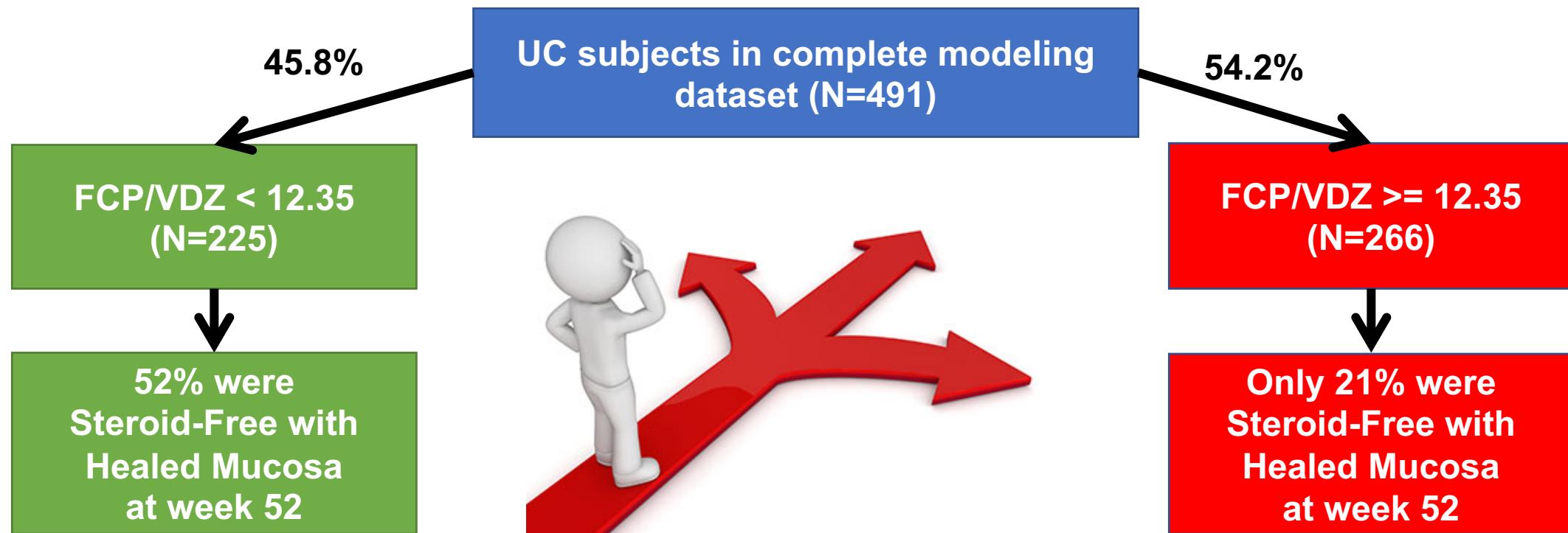
---

Week 6 Vedolizumab Level

# Results with Simple Predictor

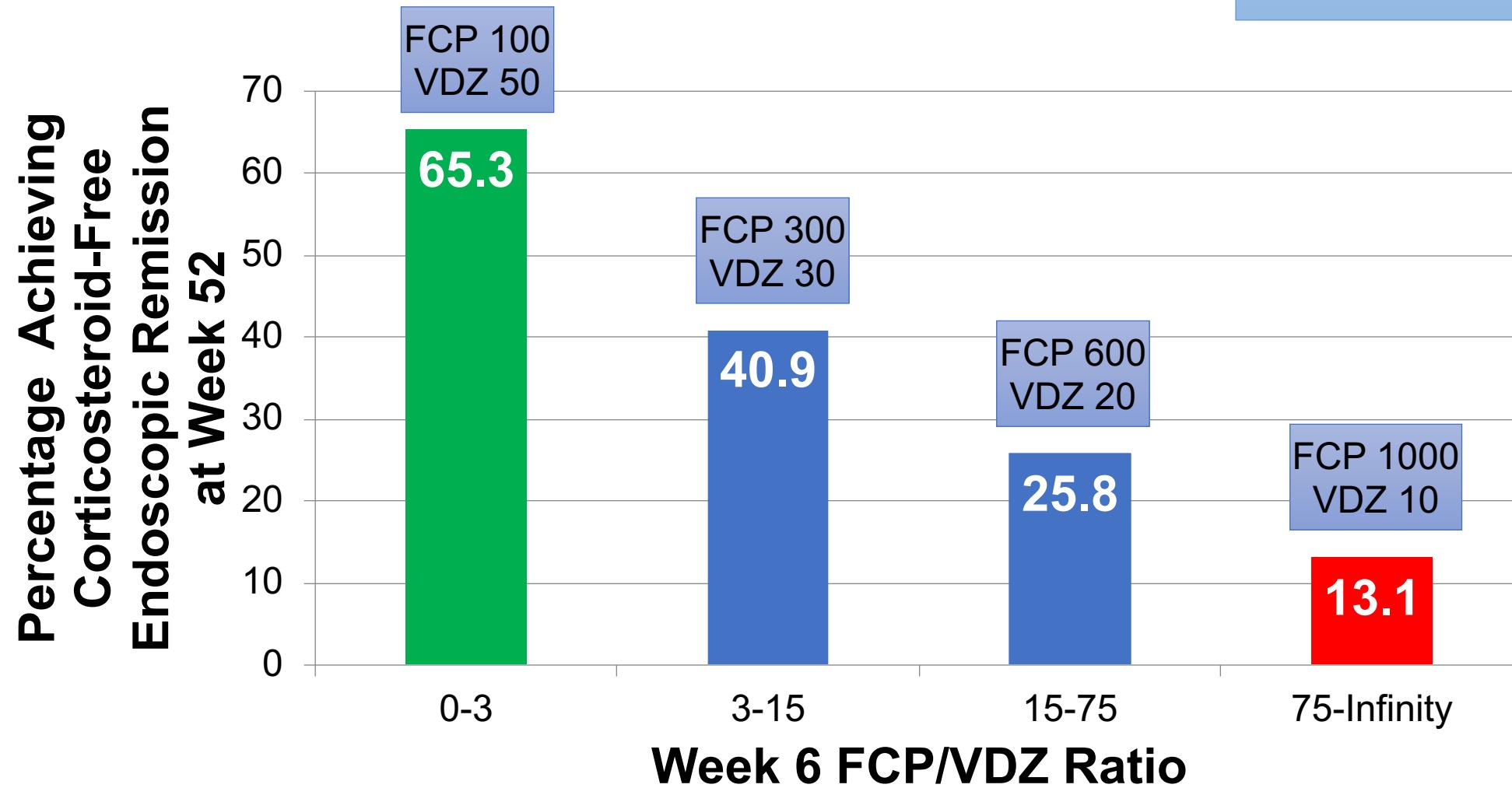
Week 6 Fecal Calprotectin  
\_\_\_\_\_  
Week 6 Vedolizumab Level

- A week 6 FCP/VDZ ratio of < 12.35 had an AuROC of 0.71
  - 95% CI: 0.67 - 0.76



# Results with Simple Model

Week 6 Fecal Calprotectin  
Week 6 Vedolizumab Level



# Conclusions

- Random forest models using data through week 6 can accurately classify UC subjects on vedolizumab as likely to succeed or fail in achieving steroid-free week 52 endoscopic remission.



# Speculations

- For patients classified as likely to fail, the optimal therapeutic approach is not clear, but could include:
  - Changing to a new medication class
  - Adding a booster induction medication (anti-TNF, JAK, anti-IL23)
  - Shortening the interval
  - Increasing the dose
- For the 53% classified as likely to fail, additional RCTs are needed to optimize the use of vedolizumab in this group of patients.

Triple combination trial: Vedo + ADA + MTX  
Co-Induction trial: IFX + Vedo

# Can We Do this In Crohn's Disease?

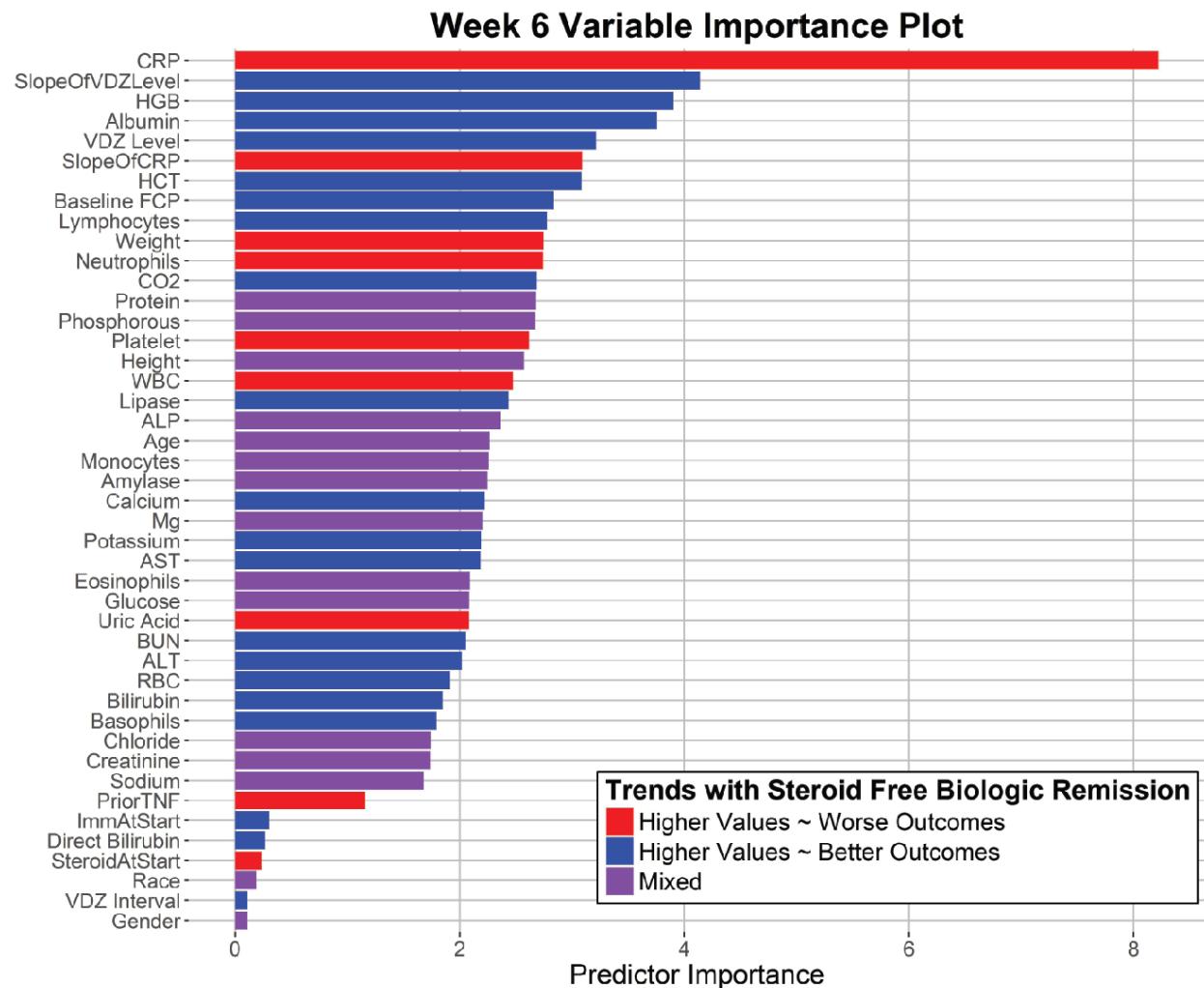
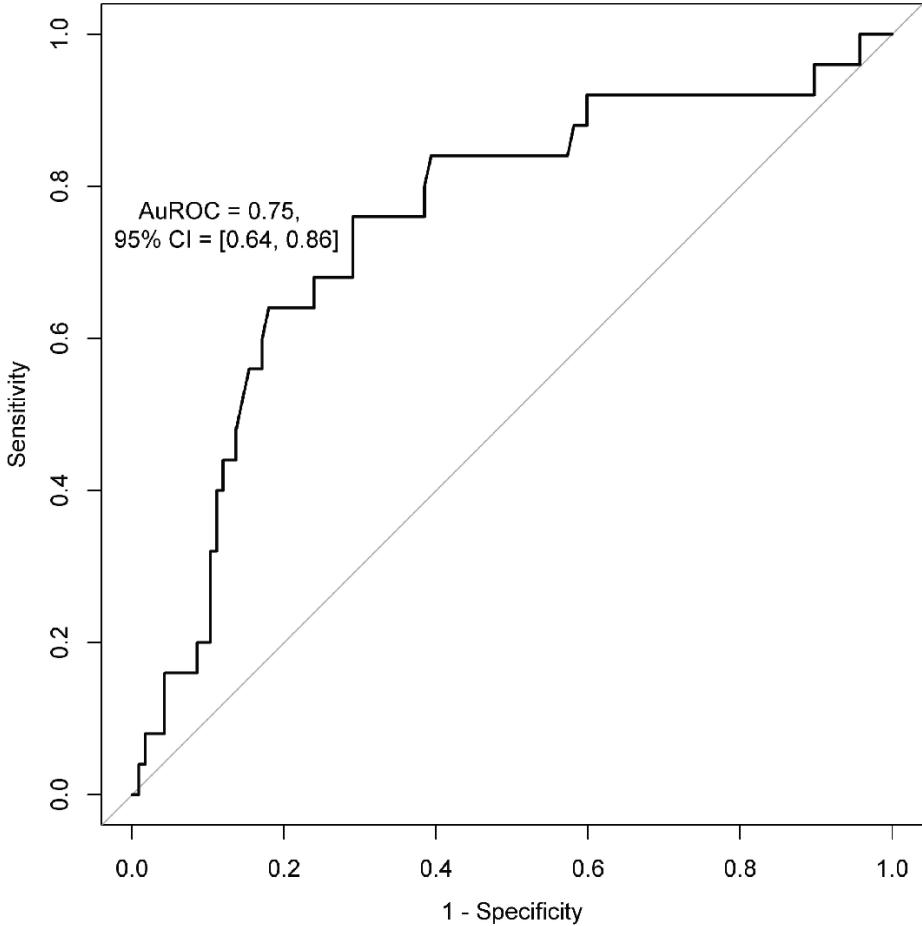
- Vedolizumab is less effective in CD
- Only 14% Clinical Remission at Week 52
- Endpoint: steroid-free biologic remission
- Can data through week 6 identify two groups of patients:
  - Those likely to succeed at a high rate
  - Those likely to fail to achieve steroid-free biologic remission



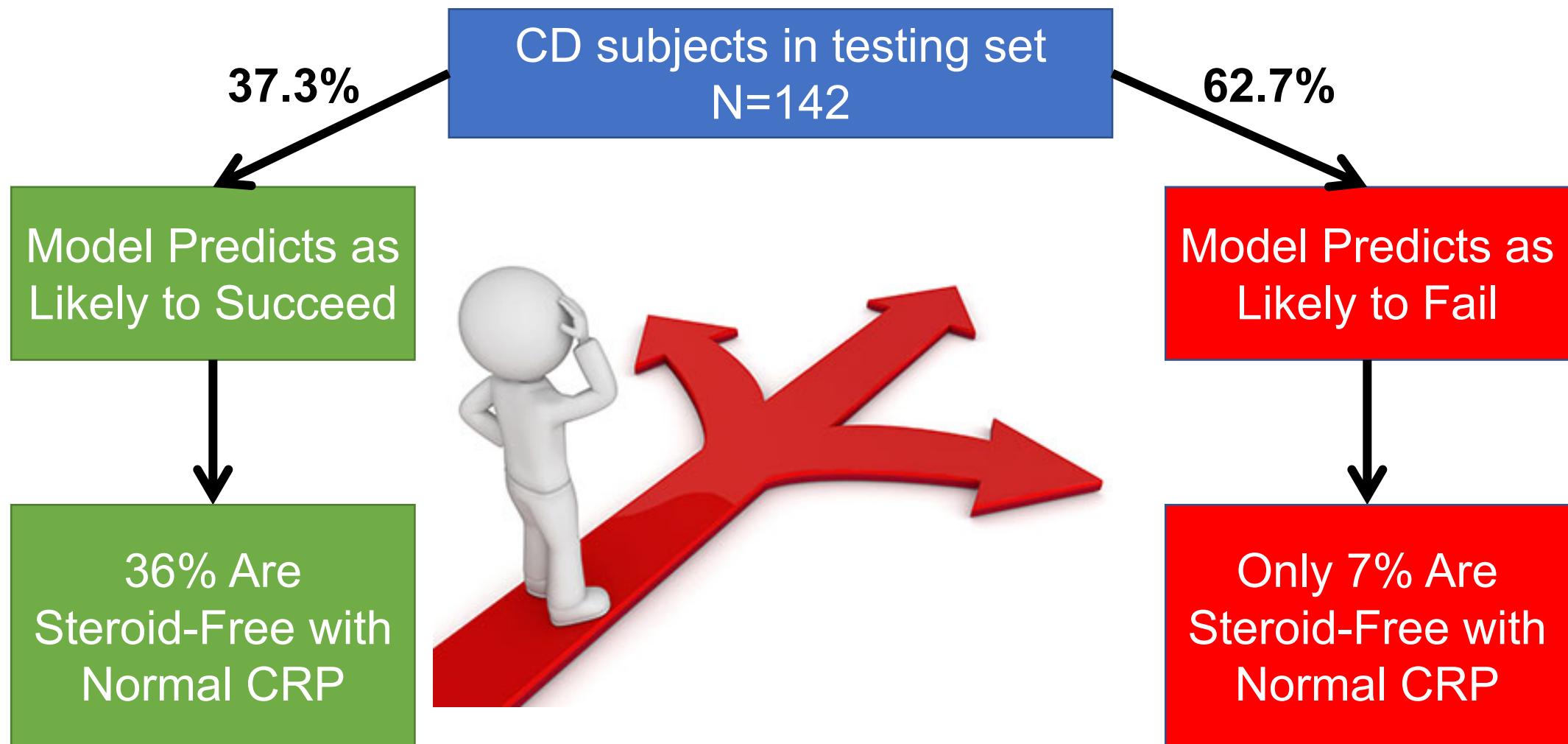
# Methods

- Exclusions: subjects randomized to placebo, OR with missing predictors or outcome data, or a ***normal CRP***
- Success endpoint: steroid-free, with normal CRP at week 52
- Randomly divided complete data into 70% model training set and a 30% model testing set
- Trained model with 1000 RandomForest trees on data:
  - Through week 6 (before 3<sup>rd</sup> dose of Vedolizumab)

# AuROC and Variable Importance for CD



# Prediction Success with Week 6 Model



# Can You Use a Simple Predictor?

- Not as simple as UC
- If Predictor > 186,  
subject is likely to achieve week 52 CS-free biologic remission
- Tested on entire (N=472) dataset,  
AuROC = 0.75 (95% CI: 0.70-0.81)

Predictor =

$$\frac{\text{Week 6 Hgb} \times \text{Alb} \times \text{Vedo Level}}{\text{Week 6 CRP} \times \text{wt (kg)}}$$

# Applications?

- Prospective validation is needed
- Use of this predictor could support clinical decisions to
  - Continue medication
  - Change to an alternative therapy at week 6
  - Increase dose or shorten interval
  - Add a booster co-therapy
  - A prospective study to compare these options is needed



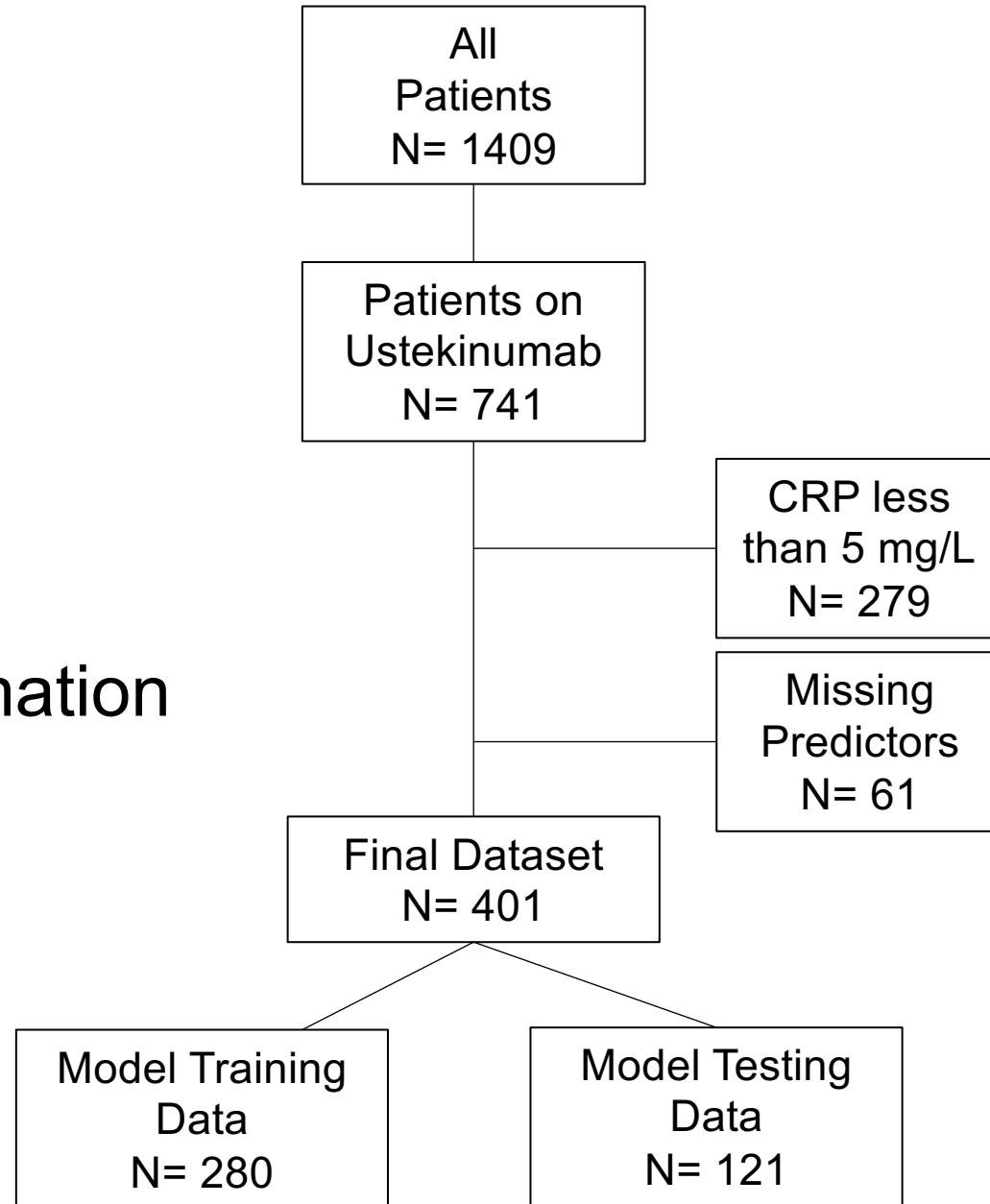


# Ustekinumab Machine Learning

Will this Patient benefit from this anti-IL12/23 Therapy?

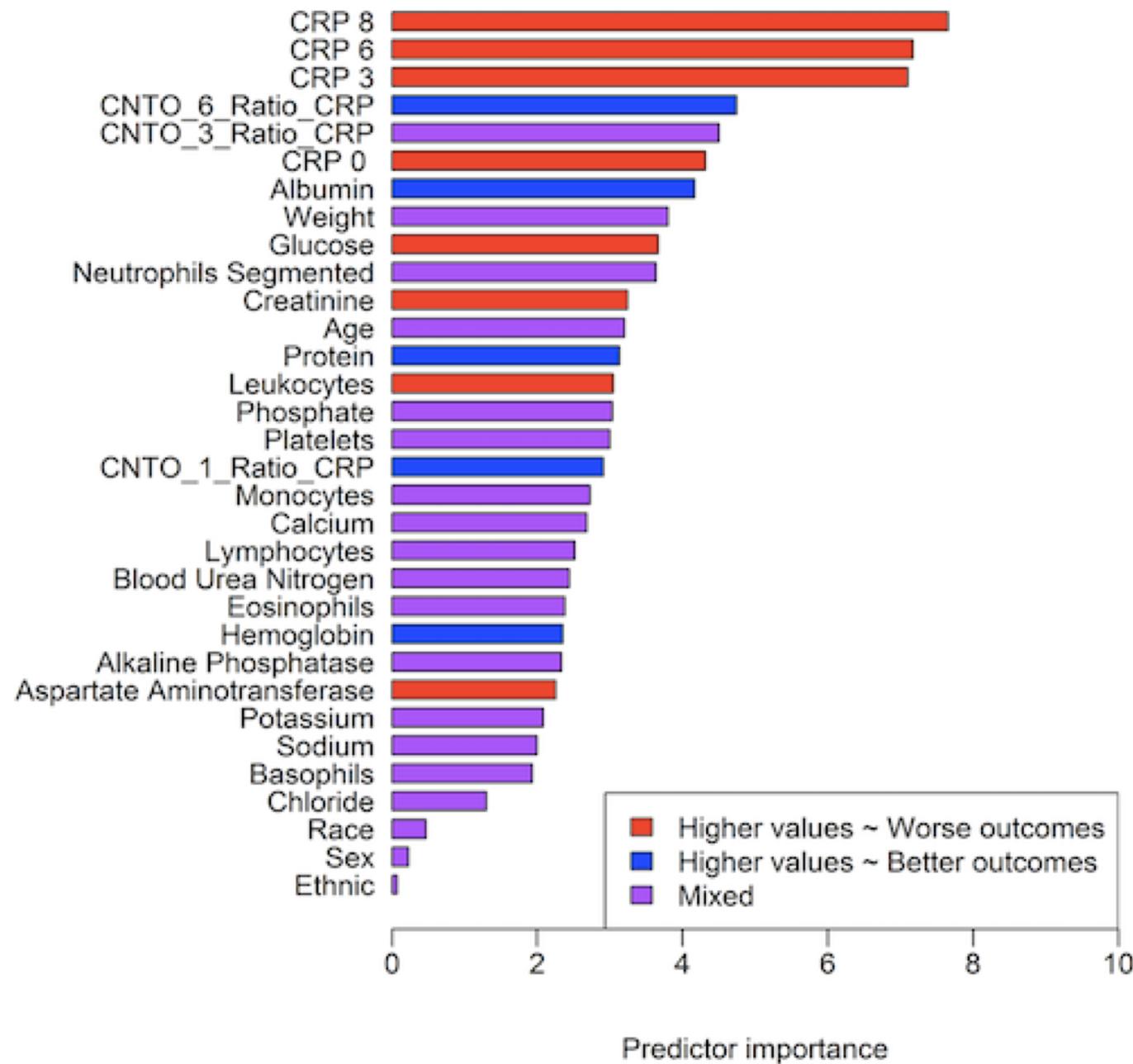
# Ustekinumab Data

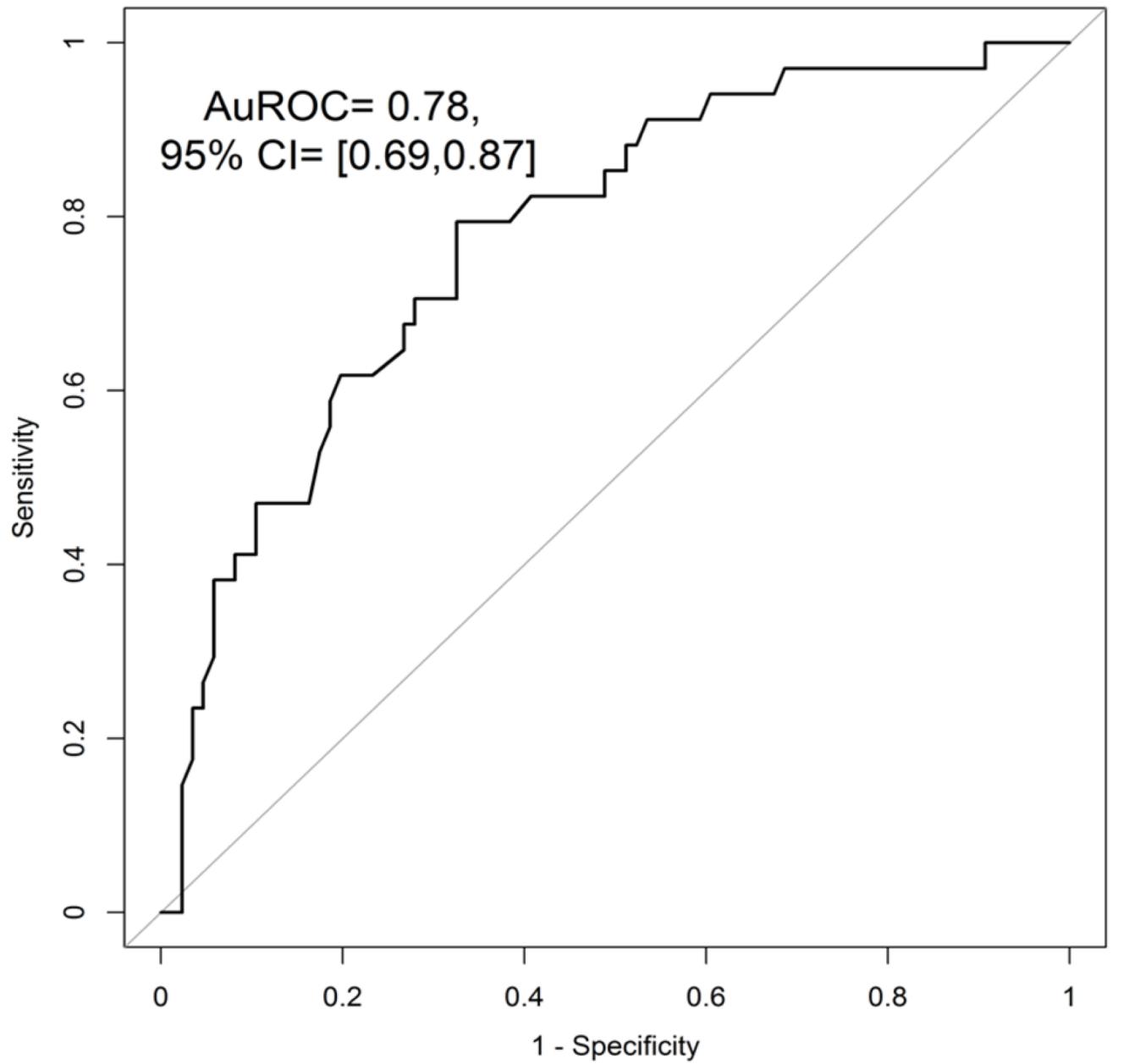
- Data from YODA
- Phase 3 trials
- Drug vs. placebo
- Required measurable inflammation
  - CRP > 5 mg/L at entry
- Outcome:
  - Steroid free remission beyond week 42



# Predictor Importance

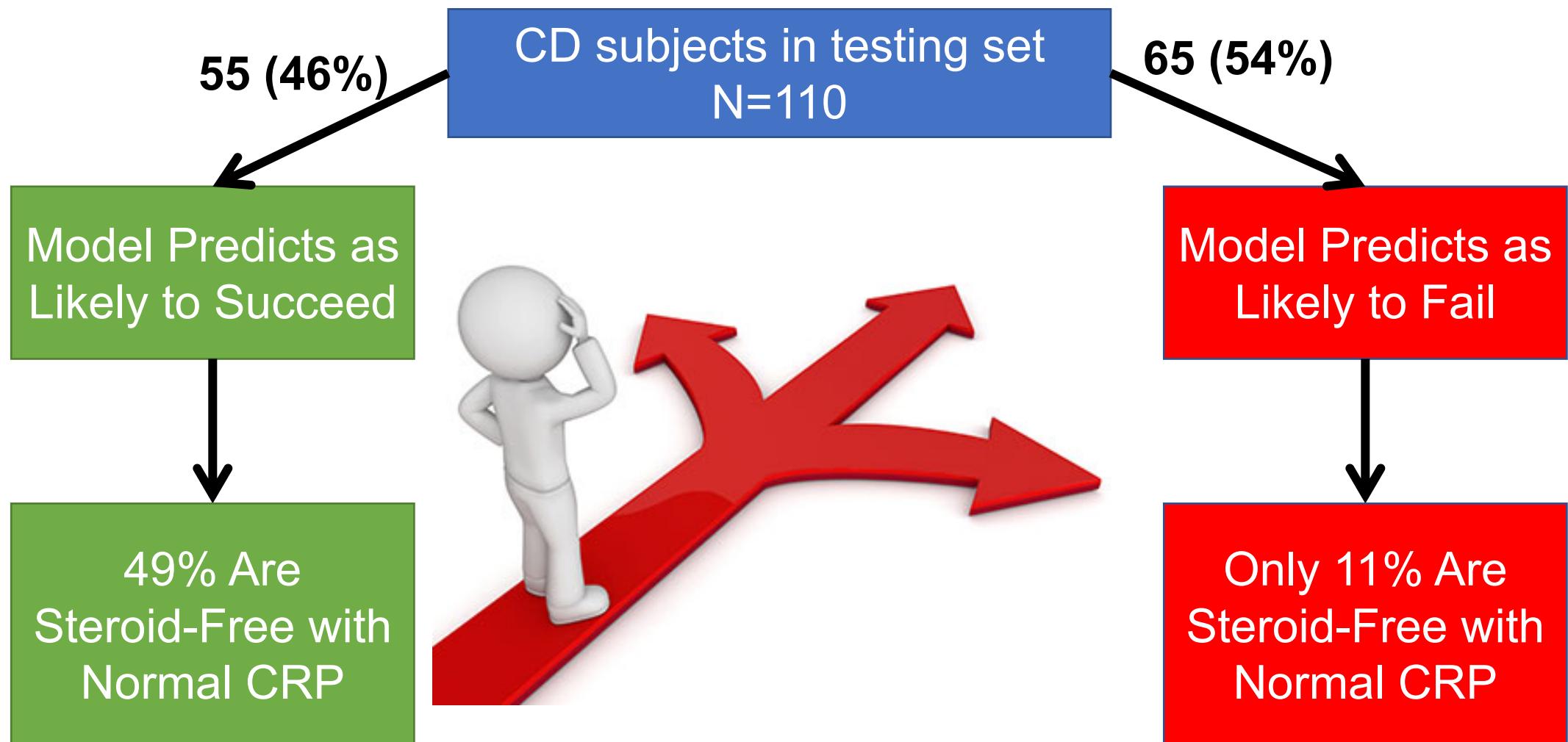
Inflammatory burden >>  
Drug levels





AuROC  
curve  
with week 8  
data

# Prediction Success with Week 8 Model



# Can You Use a Simple Predictor?

- Albumin/CRP ratio at week 6
- Time to decide whether to give next dose at week 8
- If Predictor > 4.92,  
subject is likely to achieve CS-free biologic remission > week 42
- Tested on entire (N=398) dataset,  
AuROC = 0.76 (95% CI: 0.71 - 0.82)

Predictor =

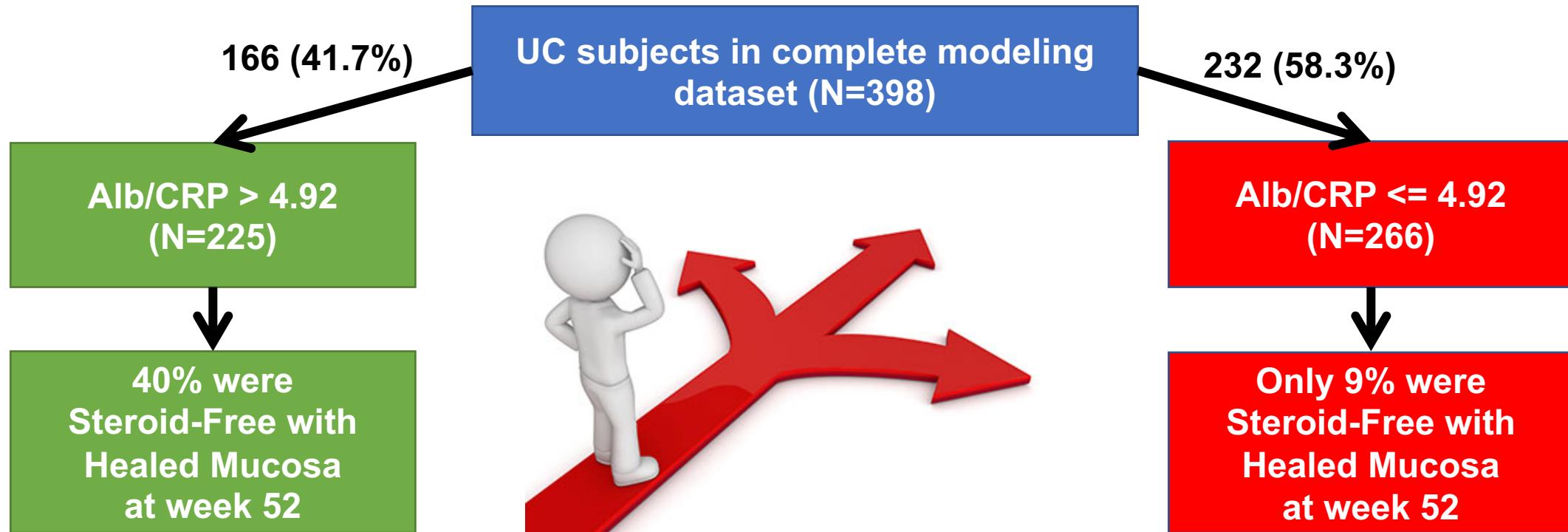
Week 6 Albumin

—  
Week 6 CRP

# Can You Use A Simple Predictor ?

$\frac{\text{Week 6 Albumin}}{\text{Week 6 CRP}}$

- A week 6 Alb/CRP ratio of  $> 4.92$  had an AuROC of 0.76
  - 95% CI: 0.71 - 0.82



# Applications?

- Prospective validation is needed
- Use of this predictor could support clinical decisions to
  - Continue medication
  - Change to an alternative therapy at week 8
  - Increase dose or shorten interval
  - Add a booster co-therapy
  - A prospective study to compare these options is needed



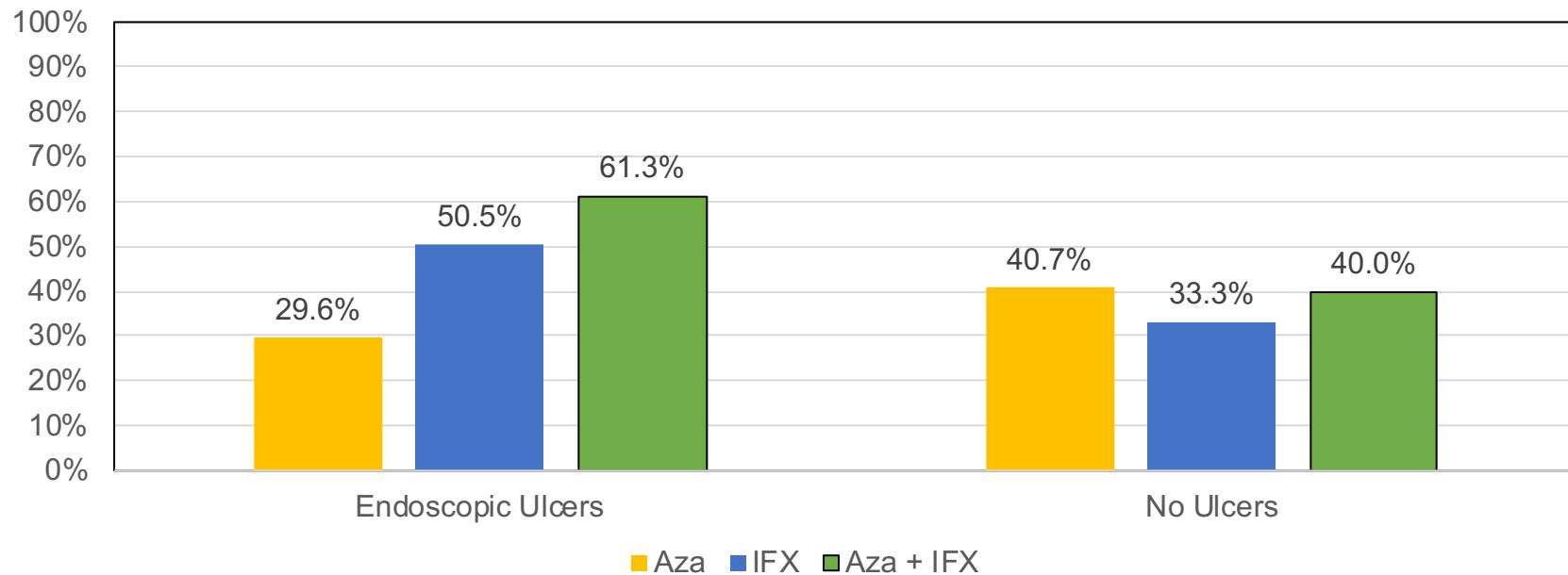
# What Have We Learned Along the Way

- We **can** predict early in the course of therapy whether a thiopurine, vedolizumab, or ustekinumab will be effective in the long term using patterns laboratory responses
- Drug levels matter a **lot** for anti-TNFs, somewhat for Vedo, but not much for Uste or Thiopurines
- Newer, better drugs are being dosed appropriately, and drug levels are becoming less important for newer biologics (and were never that important for thiopurines)

# Inflammatory Burden

- You need some inflammation for an anti-inflammatory therapy to work better than placebo

Clinical Remission At Week 26 in SONIC



# Inflammatory Burden

- We are bad at predicting long-term outcomes from lab values before starting a medication – though inflammatory burden helps
- Too much inflammation (and intestinal protein leak) make vedolizumab and ustekinumab less likely to work
  - Less important for small molecules like steroids and JAK inhibitors
- Predictors of good outcomes
  - High Albumin, Hgb, low body weight for Vedo
- Predictors of bad outcomes
  - High CRP and FCP

# Thanks To

- IBD Research Team



Akbar Waljee



Kay Sauder

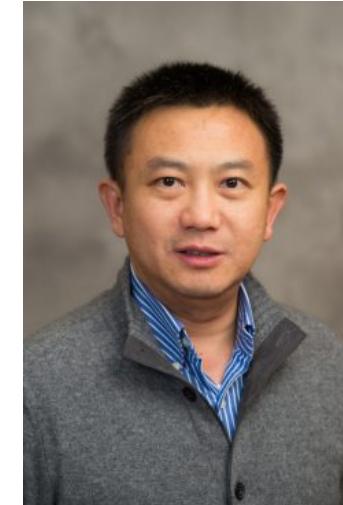


Ryan Stidham

- Machine Learning Team



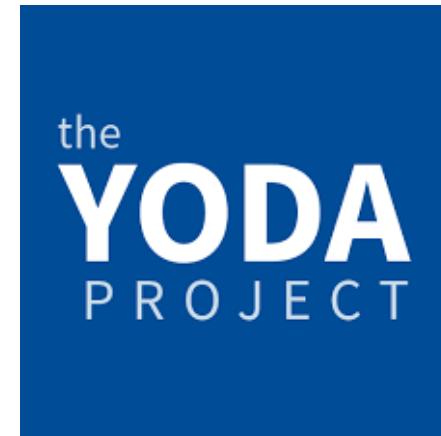
Boang Liu



Ji Zhu

# Thanks To

- CSDR – Clinical Study Data Request site  
<https://clinicalstudydatarequest.com/>
- YODA – Yale Open Data Access project
- Takeda and Janssen – for sharing their clinical trial data publicly



# Thank You

