Clover Health : Physician Efficacy Evaluation and Recommendation

Ashley Lutz

Bottom Line Up Front

- Bayesian approaches estimating provider effect on outcome with respect to covariates identified no significant provider effect
- Empirical approaches identified one provider with significantly increased failure rate.
    - Both approaches assume an NA in the outcome column is a pass to create a binary target variable.
- Given the results of both models, and the inherent randomness in the empirical approach despite bootstrapping, I do not recommend reaching out to any individual providers regarding efficacy concerns at this time.
- Additional data such as SDOH, time to follow up/event and disease characteristics would allow further analyses that both remove assumptions and add insights that may change this recommendation

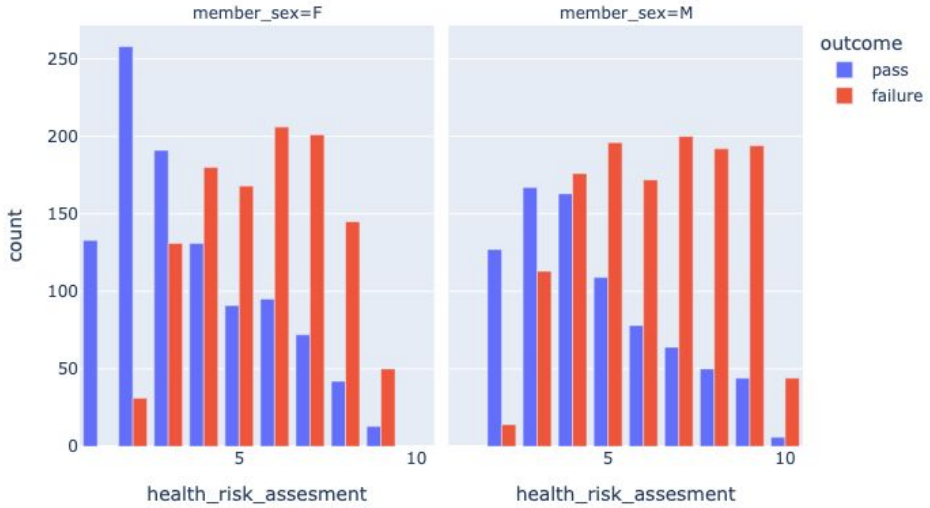
Objective

Using the data provided determine if any/all physicians are underperforming with respect to a certain condition and a certain treatment failure.

- Target variable : outcome
- Covariates : member_sex, member_age, health_risk_assesment
- Investigative variable : servicing_provider_id

Understanding the Dataset

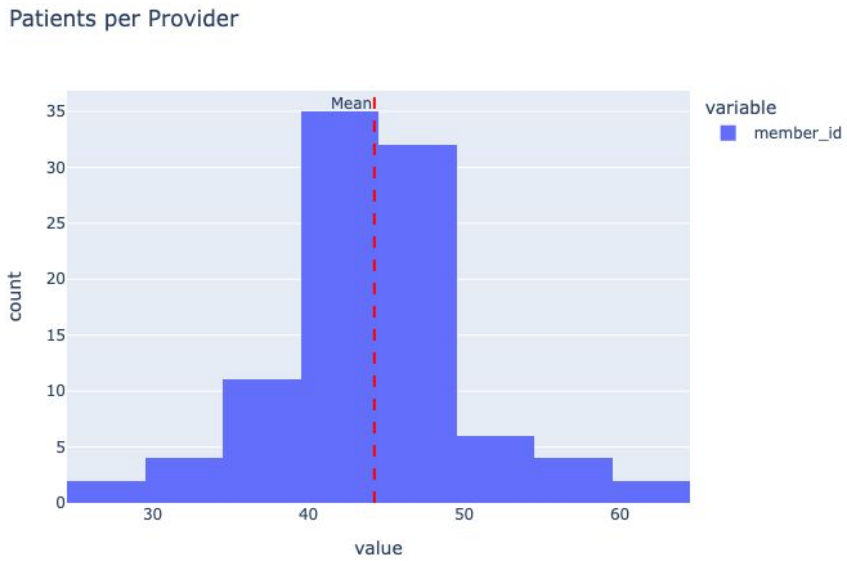
- Patients see only 1 provider
- Each patient has only 1 treatment and 1 outcome
- All providers see both genders
- All but 7 providers see patients with a range of health_risk_assesment scores

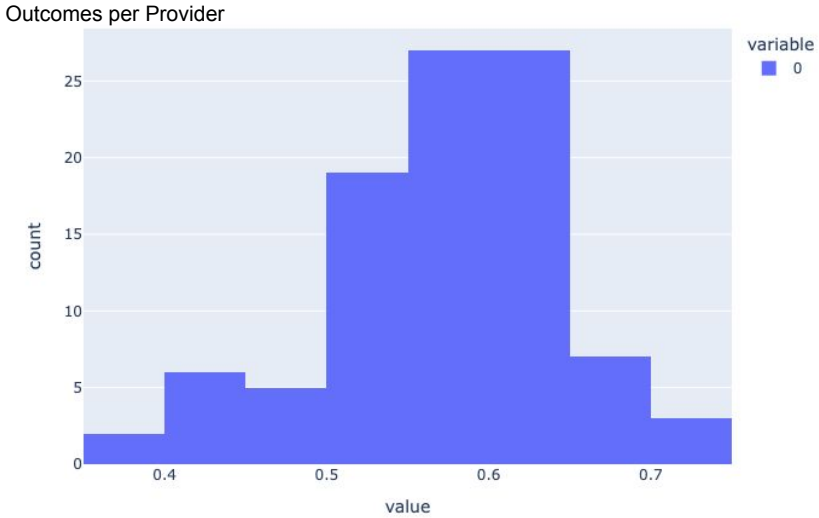


*Only female patients have health_risk_assesment scores of 1 and only male patients have health risk assessment scores of 10. This can cause issues in convergence for some models so a liblinear solver was used to compensate.

Understanding the Dataset

Providers see on average 44 patients (+-6) and the proportion of failures per provider is relatively normally distributed if not slightly left skewed.





*Failure is the only value for outcome so we **assume** an NA is a pass for now

Method 1 : Hierarchical Bayesian Logistic Regression

Answers the question: **“Does provider have an effect on outcome for a patient given their covariates?”**

Pros :

- Scalable, all providers assessed in one model
- Provides confidence interval on the certainty with which a provider is affecting an outcome while accounting for covariates

Cons :

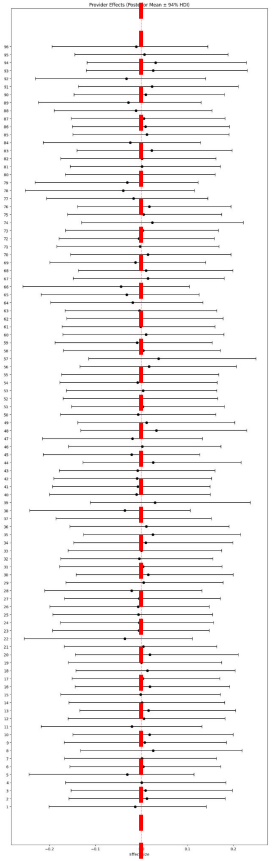
- Hard to visualize all providers at once

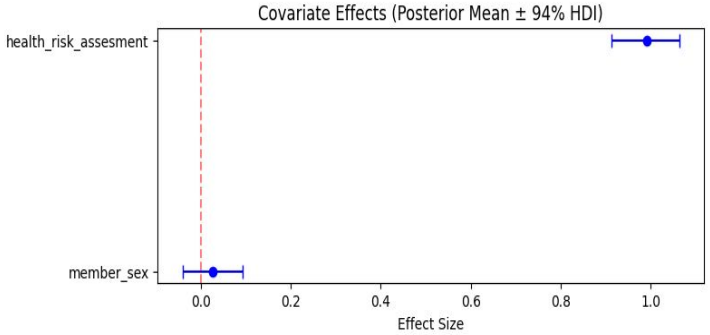
Inputs :

- Provider, member_sex, and health_risk_assesment (age removed due to collinearity)

Output :

- Outcome (Failure/pass)





Results :

- No providers have an effect with a 95% HD interval that does not cover 0 indicating little to no reliable effect.
- Health_risk_assesment score has a tight interval at average of ~1 indicating it is a main driver of outcome

Method 2 : Covariate Adjusted Empirical Null

Answers the question : **“Do any providers have a significantly different failure rate than expected given their patients covariates?”**

Pros :

- Assesses total failure rate per provider
- Providers a confidence interval for failure rates given covariates

Cons :

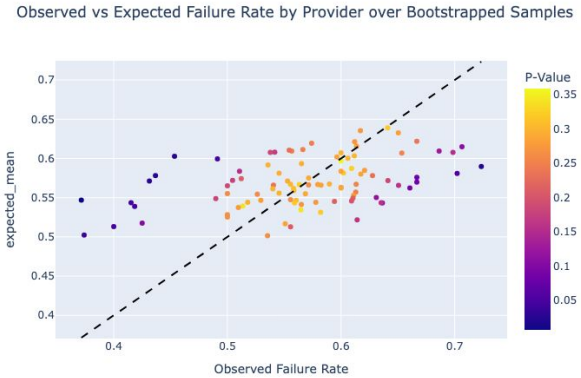
- Not scalable due to permutations and bootstrap

Inputs :

- Provider, member_sex, and health_risk_assesment (age removed due to collinearity)
- 1000 permutations for null
- 200 bootstraps

Output :

- Outcome (Failure/pass)

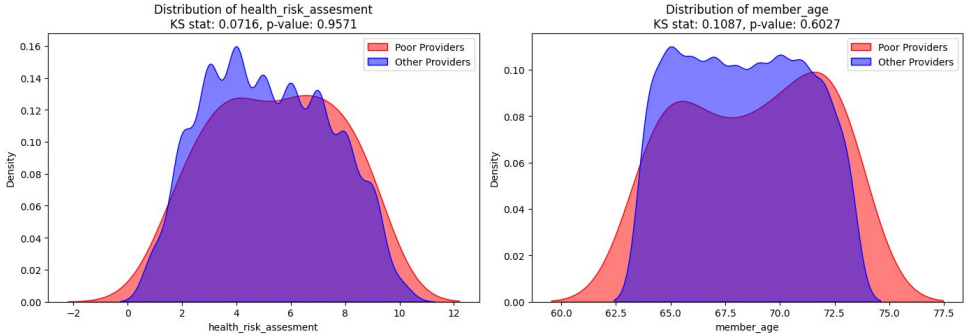


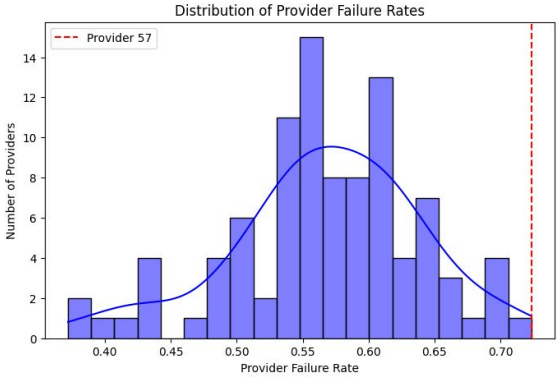
provider	observed_failure_rate	expected_mean	p_value
57	0.723404	0.589722	0.022977
38	0.491228	0.599338	0.045954
92	0.453463	0.602869	0.026973
78	0.436699	0.578083	0.026474
22	0.431373	0.571168	0.028472
5	0.418393	0.538986	0.046953
89	0.415230	0.543421	0.045954
65	0.373941	0.502332	0.041958
66	0.371460	0.546906	0.006993

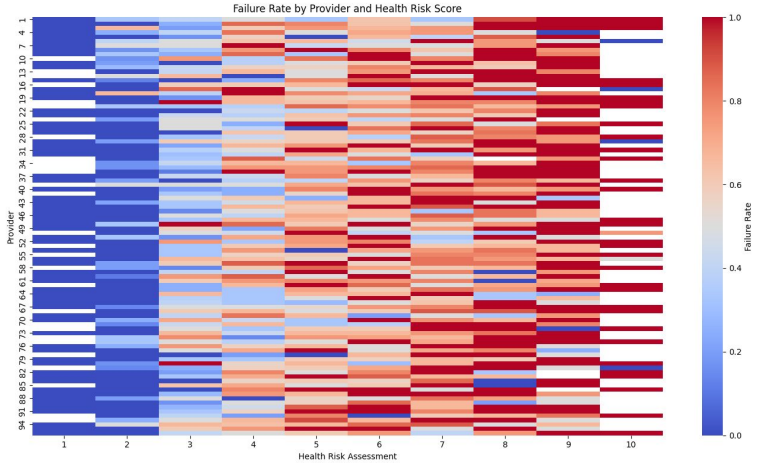
Results :

- One provider has significantly increased failure rate compare.
- Most provider actually have reduced failure rates from expected

Empirically determined poor performing provider does not differ from high performers with respect to any known covariates







Provider identified via empirical approach has the highest failure rate of the population. However at the population level it is clear that health_risk_assessment has a stronger effect on failure rate than provider, supporting the conclusions from the bayesian approach

Recommendations and Next Steps

- At this point not necessary to individually address any providers
  - While one provider did have significantly increased failure rates via empirical assessment this represents a small percentage of the total population of perspective providers and could be explained with more insight into patient characteristics, disease state, ect.
- Additional data would enable more thorough analyses that may update recommendations
  - SDOH, disease duration, contraindications, general health characteristics, and follow up/event dates would provide further insight into provider effect.
  - Specifically follow up and event dates would allow an assessment on the effect of provider on time to failure or event which may be especially insightful if this is a chronic condition.