Risk-Modeling Predictive Approaches to Treatment Effect Heterogeneity in a Randomized Trial of Epidural Steroid Injections

Presenters: Joanna, Yitao, Pinyan

Sponsors: Department of Biostatistics, Department of Radiology and Neurological Surgery, UW

March. 8th, 2021

Agenda

- Background
- Data
- Approach
- Project Study Flow & Results for aim 1
- Project Study Flow & Results for aim 2
- Conclusion
- Discussion
- Q&A

Background

Research field : Lumbar spinal stenosis;

Common treatment: Epidural steroid injections (ESIs) ;

Previous study: No short-term benefit as compared with epidural injection of lidocaine alone in a randomized clinical trial

Traditional RCT for effect modification: Subgroup analysis;

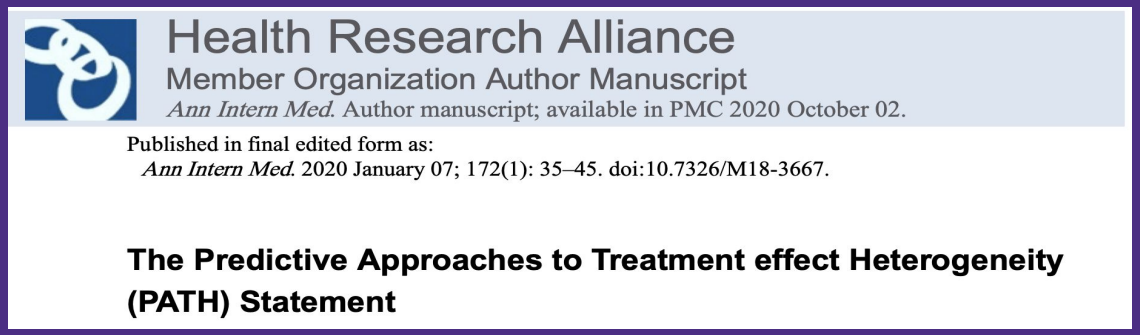
Problem: Low statistical power; Multiplicity

Background

Overall goal: Examine the ***Heterogeneity of Treatment Effect (HTE) of*** LESS trial following **PATH*** statement by taking into account all relevant patient attributes simultaneously and find out the **subgroup** that benefits from ESIs .

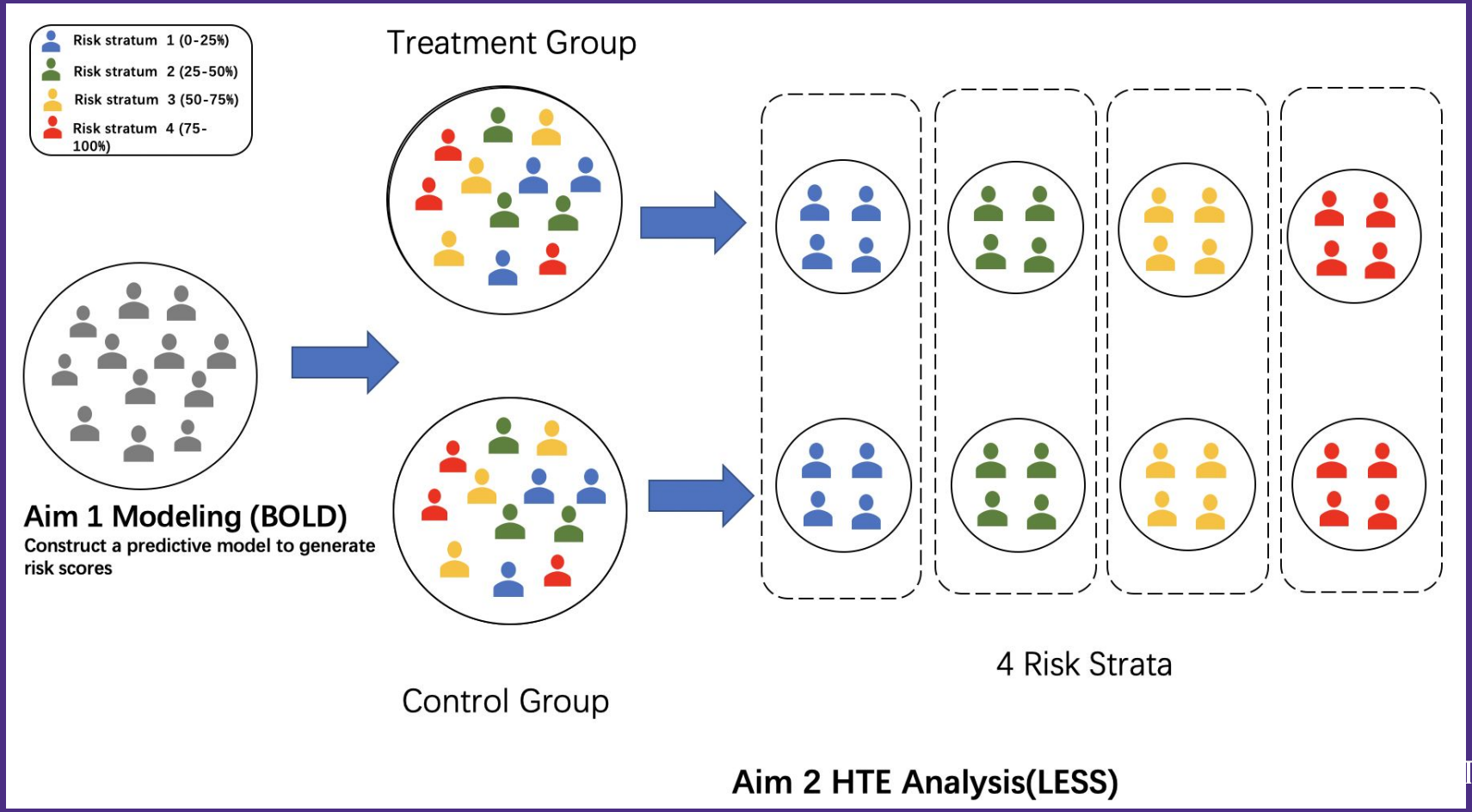
Significance: Essential for RCT in personalized medicine; provide patient-centered evidence in support of decision making.

*PATH: Predictive Approaches to Treatment effect Heterogeneity



Background-PATH

Risk modeling: a two-step modeling process, in which treatment effects are estimated in various predicted risk (categorical/continuous)



Data-description

- **BOLD**
  Back-Pain Outcomes Using Longitudinal Data (BOLD) Cohort.
  ~5200 patients, age >= 65 years from 3 integrated health care systems. [1] [Harvard Vanguard (Boston), Henry Ford Health System (Detroit), Kaiser-Permanente Northern California]

- **LESS Trial**
  A randomized, double-blinded trial.
  400 patients with lumbar central spinal stenosis
  Treatment: Epidural Steroid Injections+Lidocane
  Control: Lidocane

- **Data Type:** Electronic Health Records (EHR) and patient-reported outcomes (PROs)
  *Half of the LESS cohort (221 patients) had EHR data

1. Jarvik JG, Gold LS, Tan K, et al. Long-term outcomes of a large, prospective observational cohort of older adults with back pain.

Approach

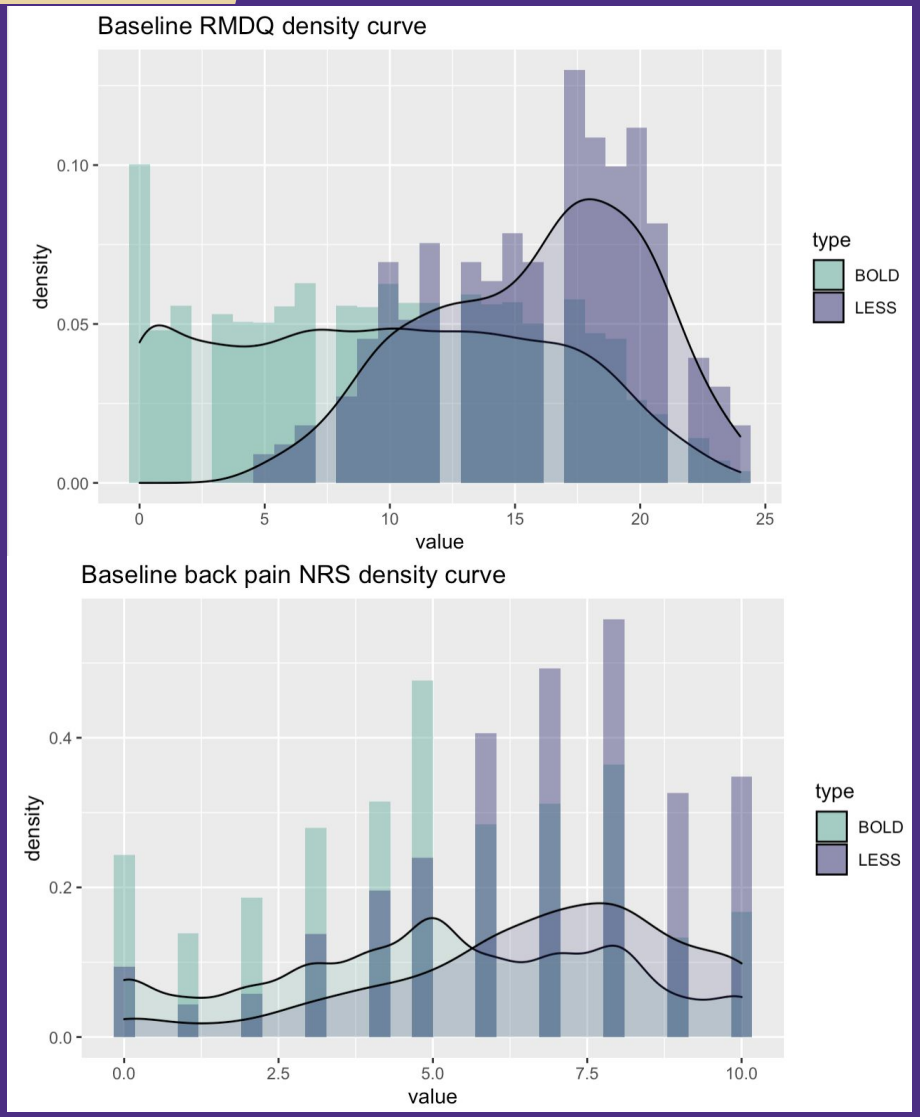
Aim 1: Developed a model to derive risk scores in BOLD cohort using LASSO

- **Primary outcome** - continuous
  3-month RMDQ risk score (measure functional limitations/disability)
- **Secondary outcome** - continuous
  3-month back pain NRS risk score (measure pain)

Aim 2: Evaluated heterogeneity of treatment effect

Step 1: Applied the model to predict risk scores within the LESS trial patients.
Step 2: Examined risk-based variation (categorical/continuous) in treatment effects.

Distribution of outcomes



Patients in the LESS cohort had more severe disability and pain intensity than the BOLD cohort

Exclusion criteria: Excluded patients with baseline RMDQ <= 2 or back pain NRS scores <= 1 to adjust for inconsistency

Project Study Flow-Aim 1

Statistical analysis (Aim 1)

1. **Developed a series of LASSO regression models** to predict the risk of the outcome. 5-fold cross-validation was used to select optimal parameters.

2. **Testing and evaluation** based on R-square and MSE

Project Study Flow-Aim 1

Statistical analysis (Aim 1) - Model selection

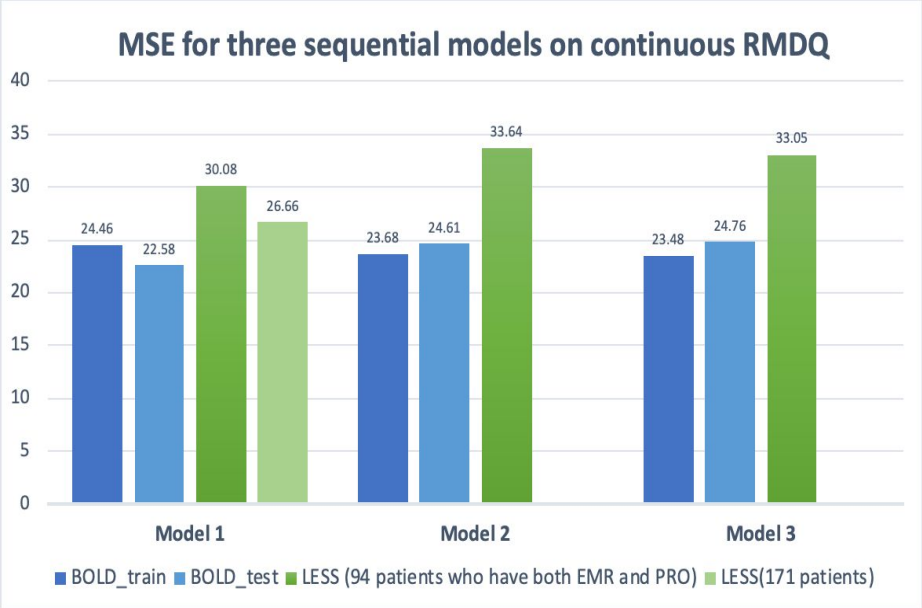
more patients to fit

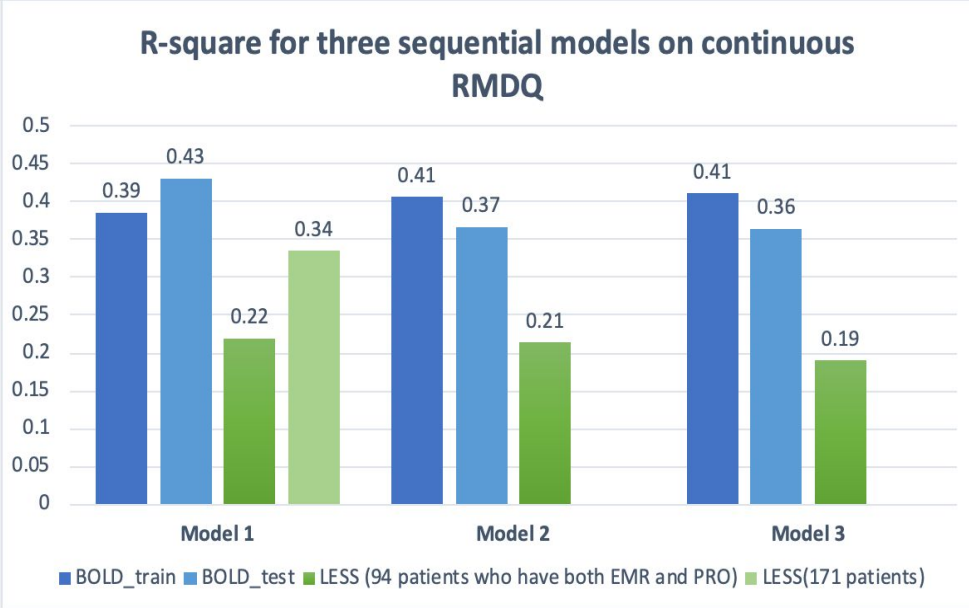
more variables selected

Model	Variables	Data Source	Population-testing in LESS with standard of care
1	age, sex, hispanic, education, marital, employment, lawyer, smokingstatus, falls, bpi, painexpect3mo, roland_baseline, painnrs_back_baseline, painnrs_leg_baseline, PHQ-4, BMI, EQ_VAS_baseline, EQ_index_baseline	PRO	Patients who have PRO
			Patients have EMR and PRO
2	Model 1 variables + study_site, ICD-9 categories , comorbidity categories, CPT categories (manual, spine-image, percut, spine-surgery)	PRO+EMR	Patients who have EMR and PRO
3	Model 1 variables + study_site, ICD-9 categories, refined comorbidity categories, refined CPT categories	PRO+EMR	Patients who have EMR and PRO

1.trade off between performance and broad applicability
2.Test Model 1 on both LESS-full and LESS-subset to check prediction performance.

Results-Aim 1 Continuous RMDQ scores

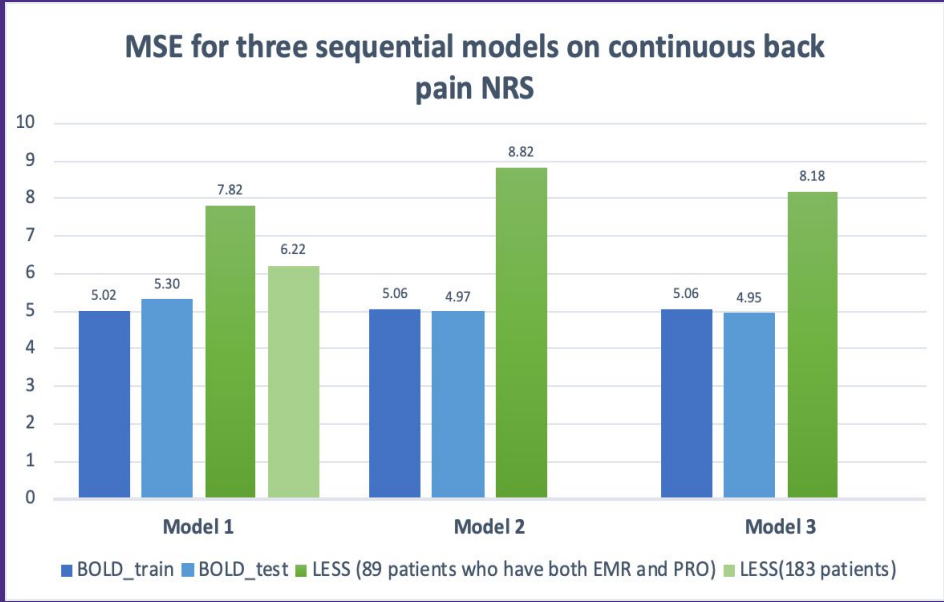


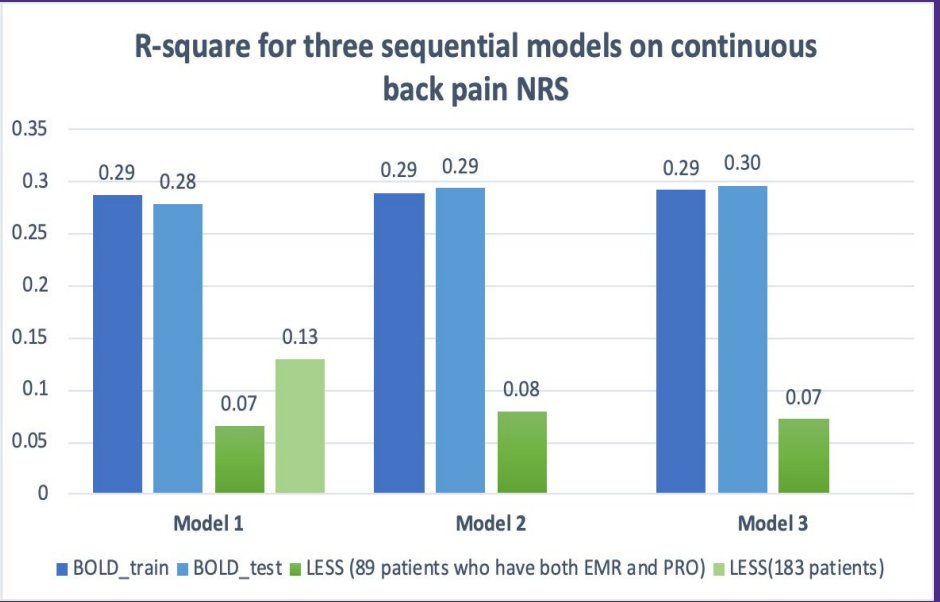


Model 1 only using variables from PRO data performed better on LESS cohort with lower MSE and higher R-square, in particular better in LESS with only PRO variables (171 patients)

Testing on LESS in control arm:
94 LESS: both EMR and PRO
171 LESS: have PRO

Results-Aim 1 Continuous NRS scores





- There weren't big differences among these three models especially when using R-square as selection criterion
- Not a good choice to derive prognostic scores based on continuous NRS scores (lower R-square compared with continuous RMDQ.

Testing on LESS in control arm:
89 LESS: both EMR and PRO
183 LESS: have PRO

Project Study Flow- Aim 2

Statistical analysis (Aim 2)

1. **Risk score calculated in the whole LESS:** calculated the risk score using **continuous RMDQ** as the endpoint with the same model coefficients generated in Aim 1

2. Divided patients into 4 risk strata and evaluate treatment effects within each strata

- Model:

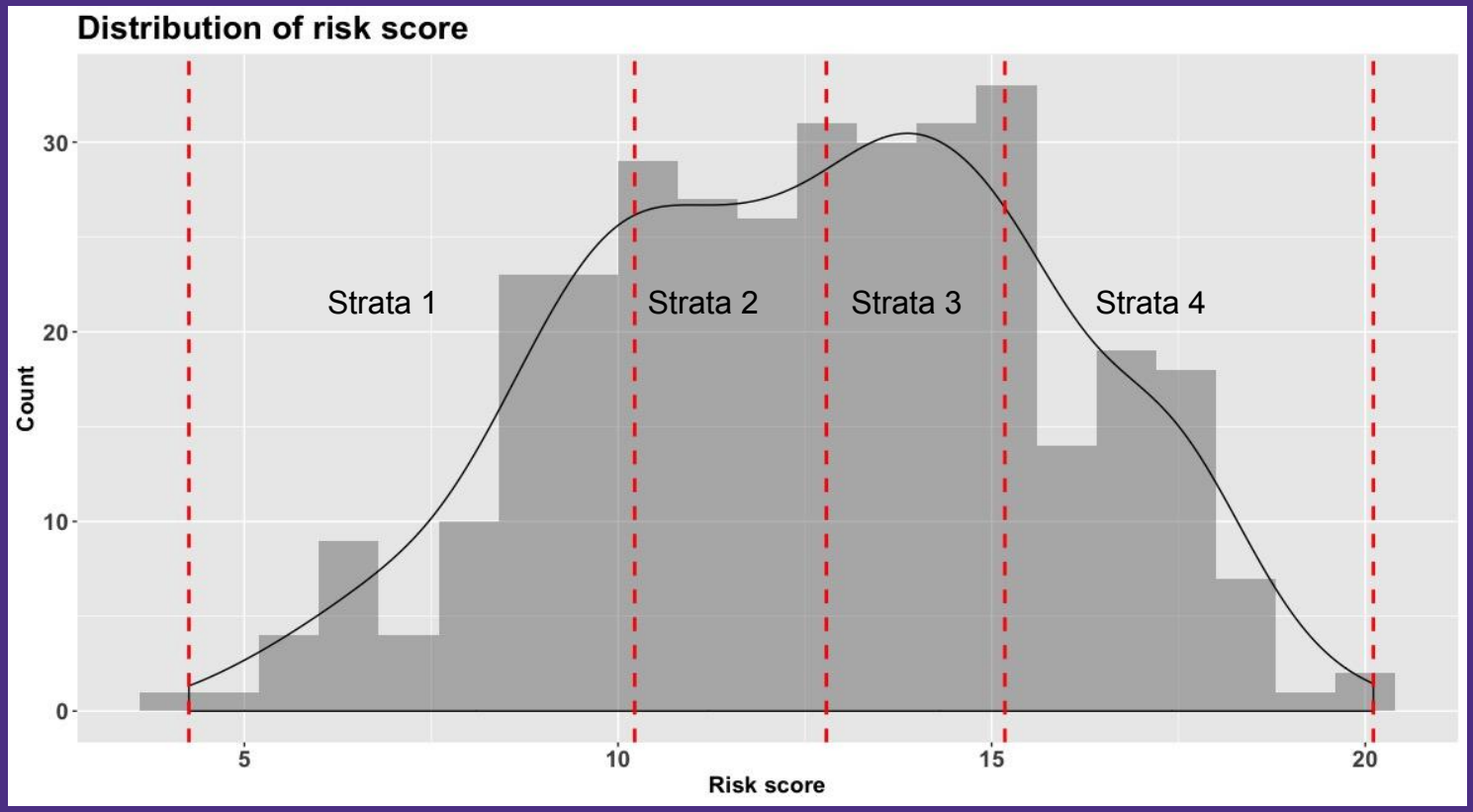
  Measured 6-week RMDQ~ strata + treatment indicator + strata * treatment indicator + covariates at baseline (study site, baseline RMDQ)

3. Estimated heterogeneity of overall treatment effects using continuous risk score (primary analysis)

- Model:

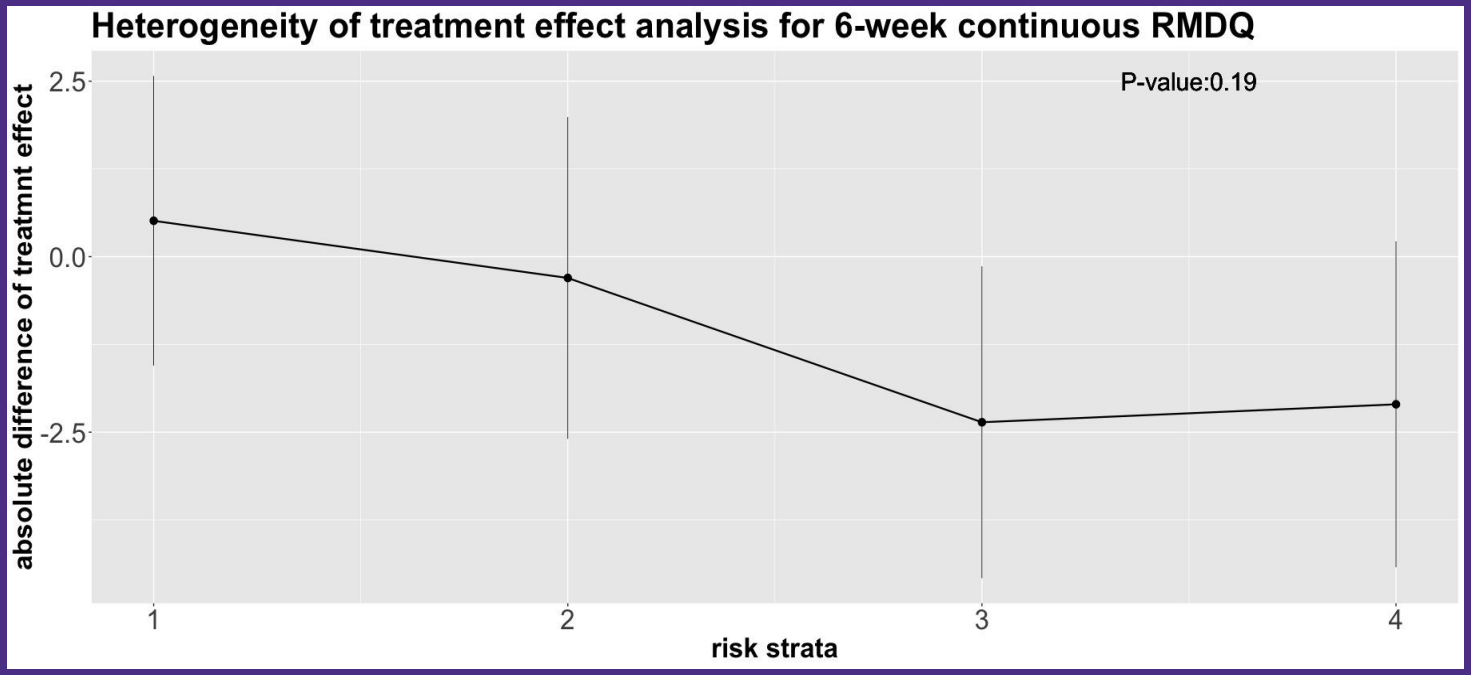
  Measured 6-week RMDQ ~ continuous risk score + treatment indicator + continuous risk score * treatment indicator + covariates at baseline (study site, baseline RMDQ)

Results - Aim2- Strata



We divided patients into quartiles based on the risk scores. For each strata from lowest risk quartile to the highest risk quartile, the range of their prognostic scores was (4.26, 10.22), (10.23, 12.79), (12.81, 15.19), (15.20, 20.11).

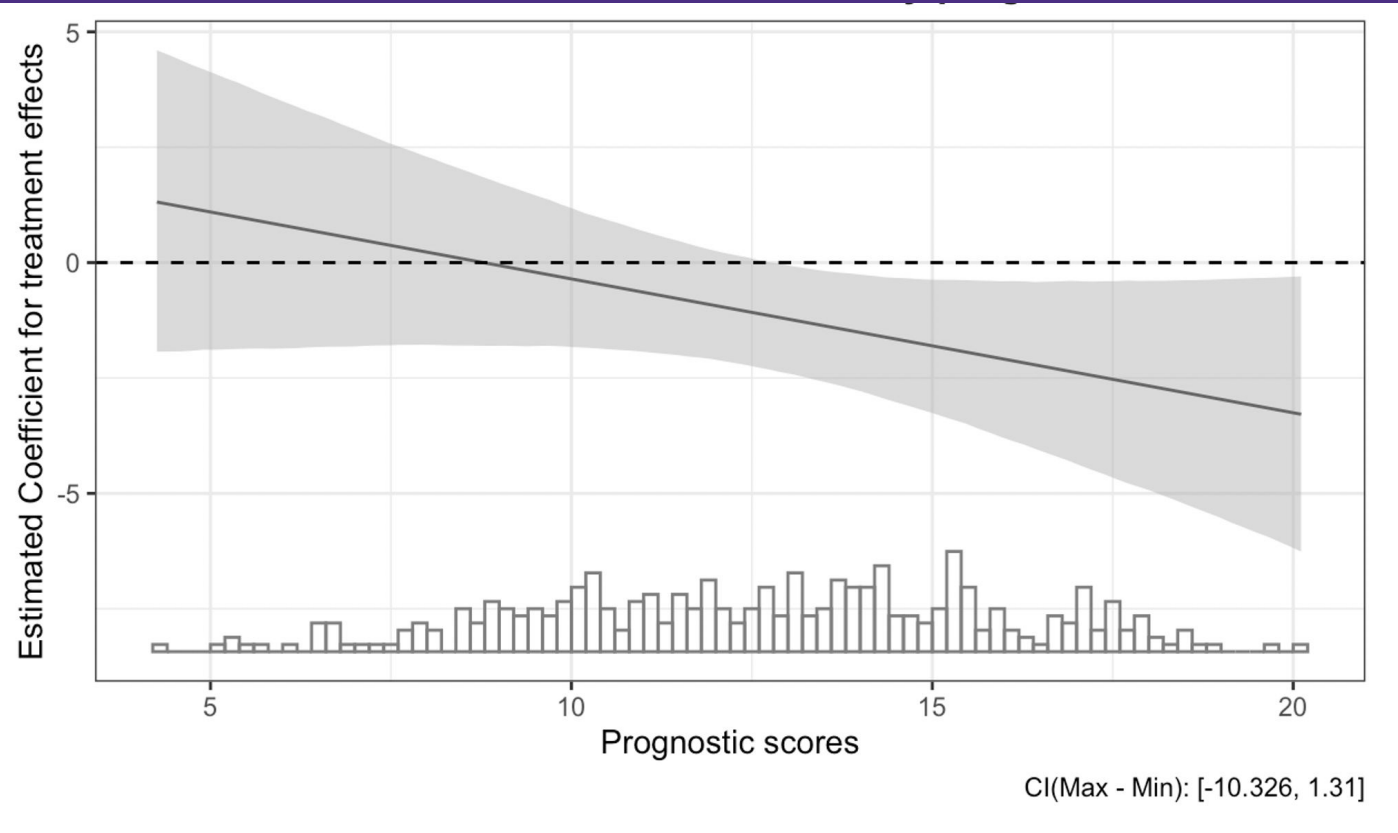
Results - Aim2 - Strata



No statistically significant difference (p-value=0.19 > 0.05)

- RMDQ: Decreasing trend from strata 1 to strata 3: patients with higher risk received more benefit from the epidural steroid injections.
- Significant treatment benefit in Strata 3 (average treatment effect, -2.36 (-4.58, -0.14)
- Similar results when using 6-week back pain NRS as the outcome

Results - Aim2 - Continuous risk score



No statistically significant difference (p-value=0.11 > 0.05) of overall treatment effect

Conclusion

1. Used a relatively novel risk-modeling approach to evaluate treatment effects

2. Heterogeneity of overall treatment effects was not statistically significant (6-week RMDQ / NRS scores)

3. The HTE analysis provided valuable insights to assess the effectiveness of epidural steroid injections.

- Discovered one strata who had significant treatment benefit

Discussion

1. **Different follow-up endpoints in Aim1 and Aim2** used 3-month RMDQ in Aim 1 and 6-week RMDQ in Aim 2 (lacking the data of 6-week RMDQ in the BOLD cohort).

2. **Only considered linear model (LASSO).** Sensitivity analysis by adding the interaction term / splines. R-squared did not improve. Splines even worse.

3. **Variables were not completely consistent in two cohorts** Trade-off between observations and variables we had to take into account.

4. **Differences between BOLD and LESS cohorts** Patients in the LESS trial had more severe disability and pain intensity than the BOLD cohort

Acknowledgements

Jerry Jarvik, Pradeep Suri, Laurie Gold, Patrick Heagerty and other BOLD research team members

Adam Szpiro and Brian Leroux, Course instructor

Q&A

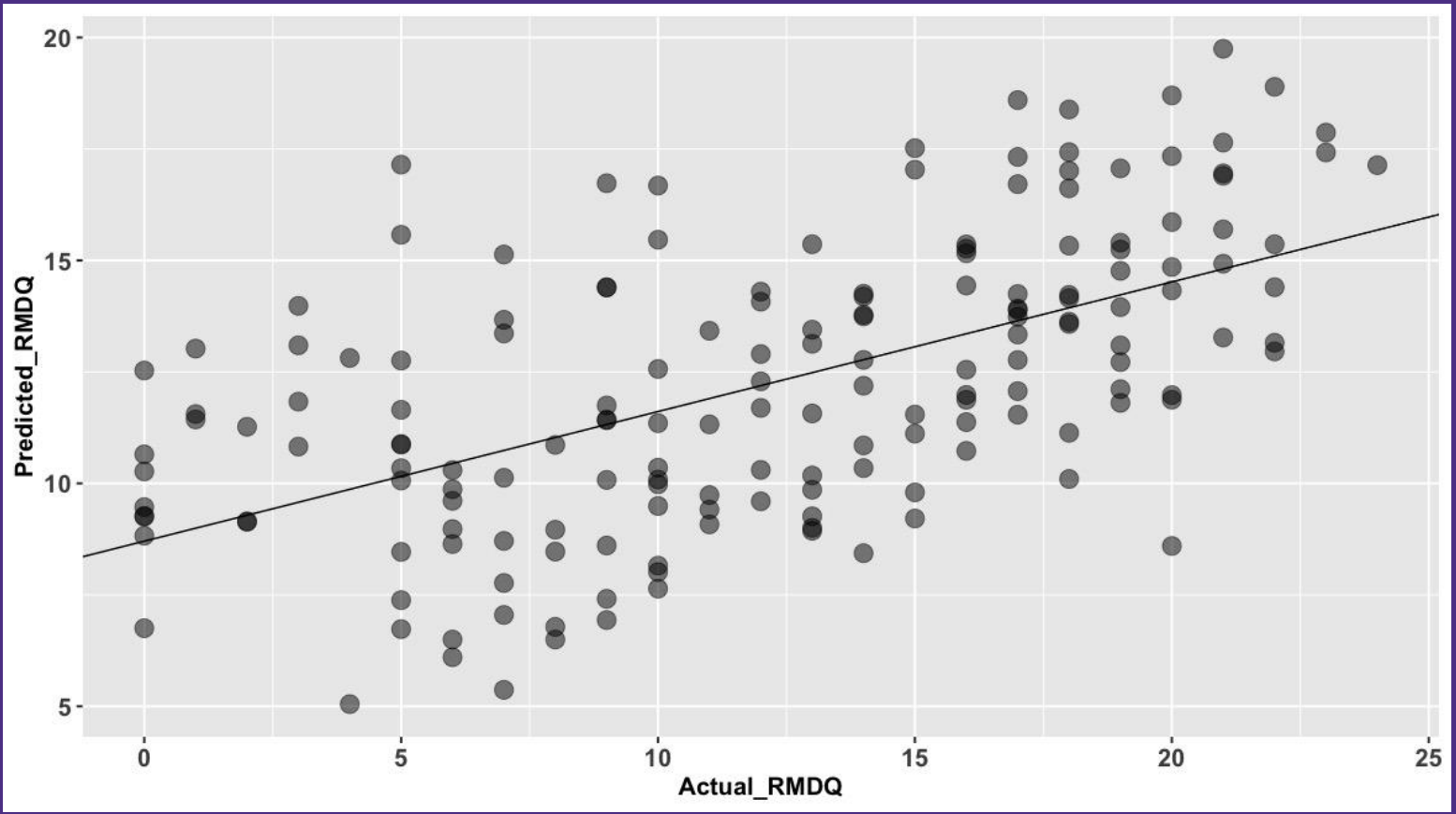
Thanks for listening!

Contribution

Role descriptions of team members

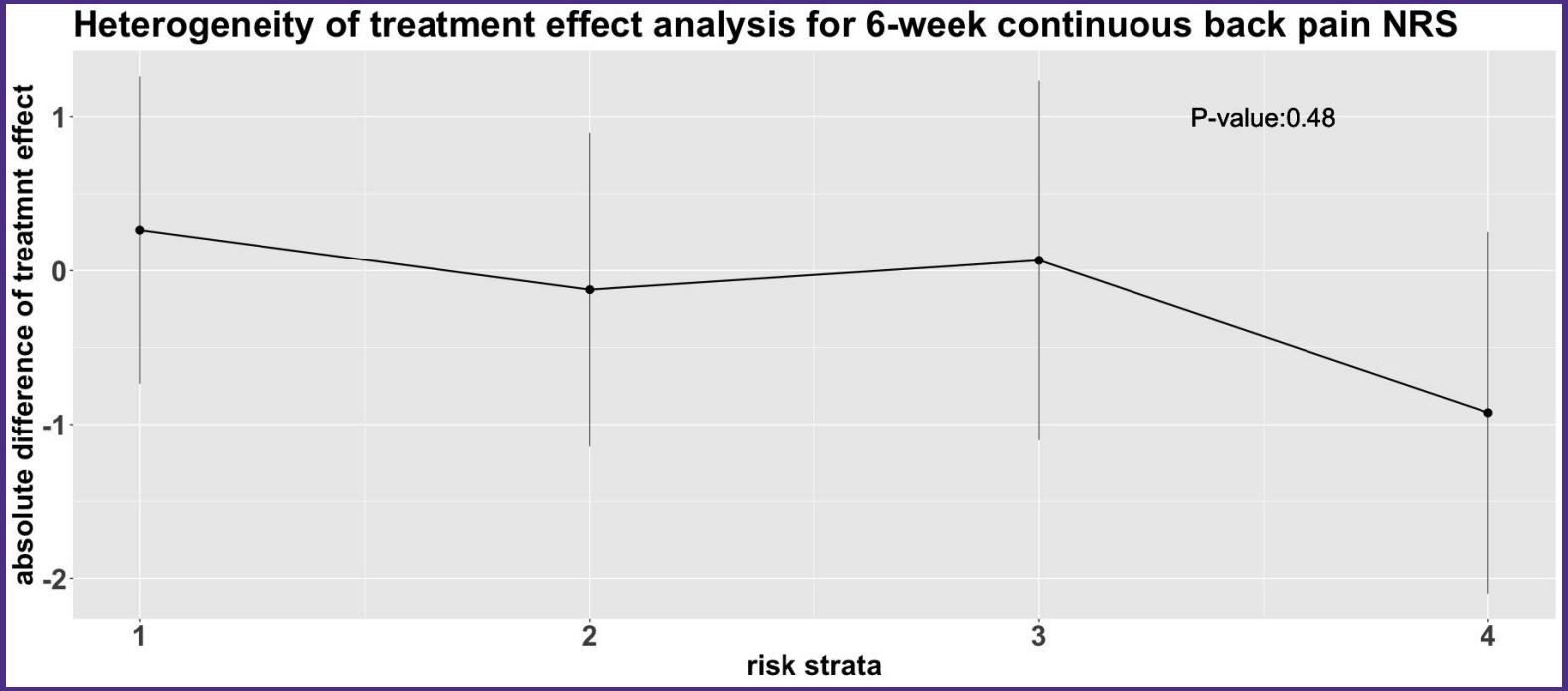
- **Group work:** analysis of continuous RMDQ; introduction; methods; discussion
- **Yitao Wu:** analysis of continuous back pain NRS continuous; scatter plots; plots of HTE analysis; interaction & spline analysis
- **Joanna Xiao:** analysis of binary RMDQ; data cleaning and merging; variables transformation and categorization; histogram; Table 5
- **Pinyan Liu:** analysis of binary back pain NRS; GAM analysis; Table 1 & Table 2; continuous interaction analysis; forest plot

Appendix Continuous RMDQ scores in Aim 1



Scatter plot of Model 1 (actual RMDQ versus predicted RMDQ) on LESS cohorts with PRO (171 patients) with the best-fitted regression line

Results - Aim2 - NRS



Same results as RMDQ.

Conclusion: No statistically significant heterogeneity of overall treatment effect in the **functional limitations /disabilities and back pain.**