P9185 - Project 3: Protocol of a Cluster-randomized trial for Asthma-PASS

Amy Pitts, Yuqi Miao

2022-04-25

Background

Our interest is in persistent asthma in minority children.

- Comprehensive school-based interventions in collaboration with communities to reduce asthma morbidity and promote physical activity in urban areas.
- A pilot cluster RCT was conducted exploring this intervention in Bronx elementary schools
  - **Goal:** whether Children in schools receiving Asthma-PASS intervention may experience a greater improvement in the number of SFD at 6 **or** 12 months follow up than the children in the comparison group.
  - 4 Bronx elementary schools were recruited into the pilot study.
  - A total of 108 children recruited including ages 4-11 years with physician-diagnosed persistent or uncontrolled asthma attending kindergarten to 5th grades

Overview

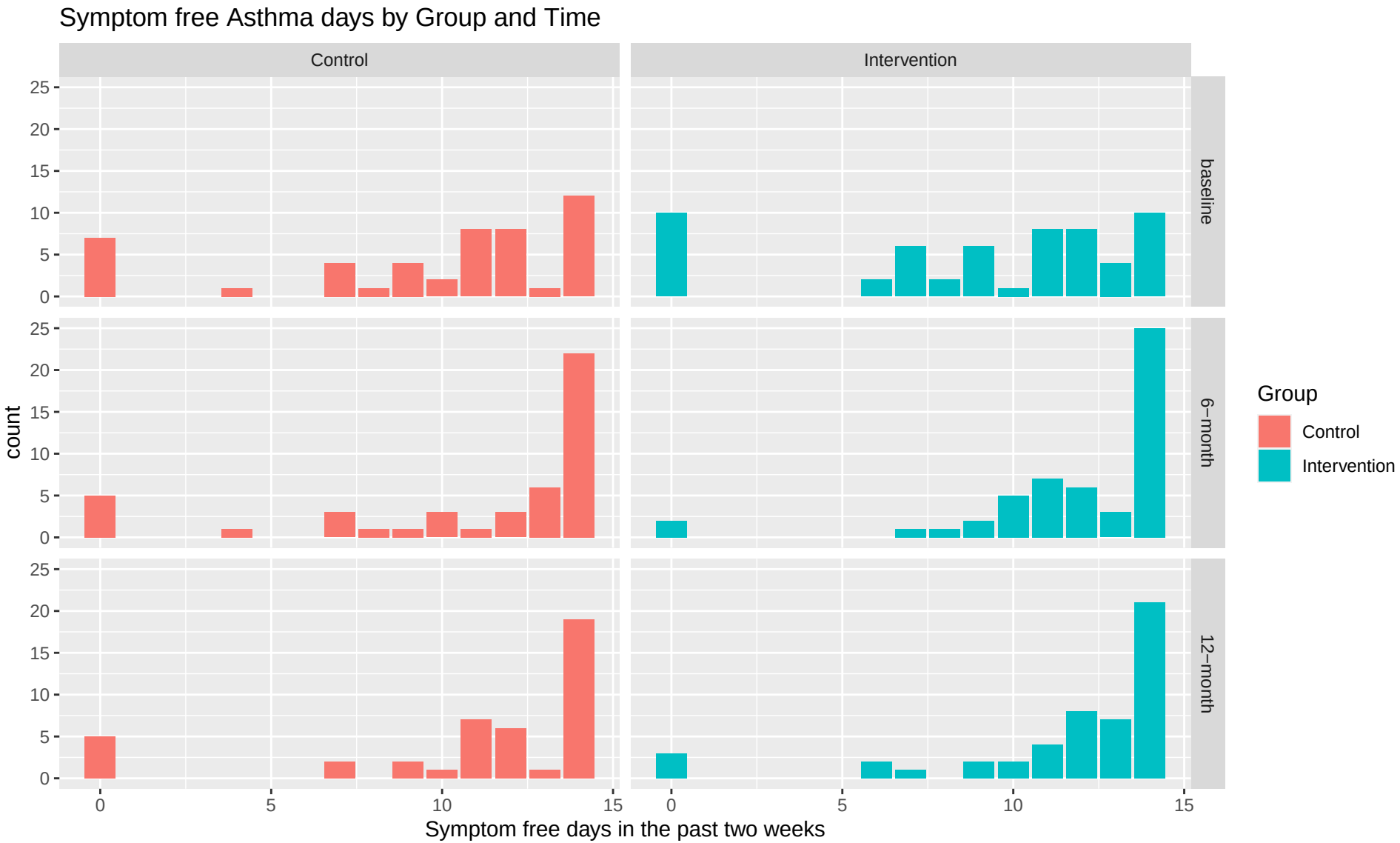
1. Data Overview
2. Exploration into Pilot Study data
   - Model Specifications
   - Results
3. Phase III Proposal
   - Model Specifications
   - Sample Size Suggestions

Data Overview

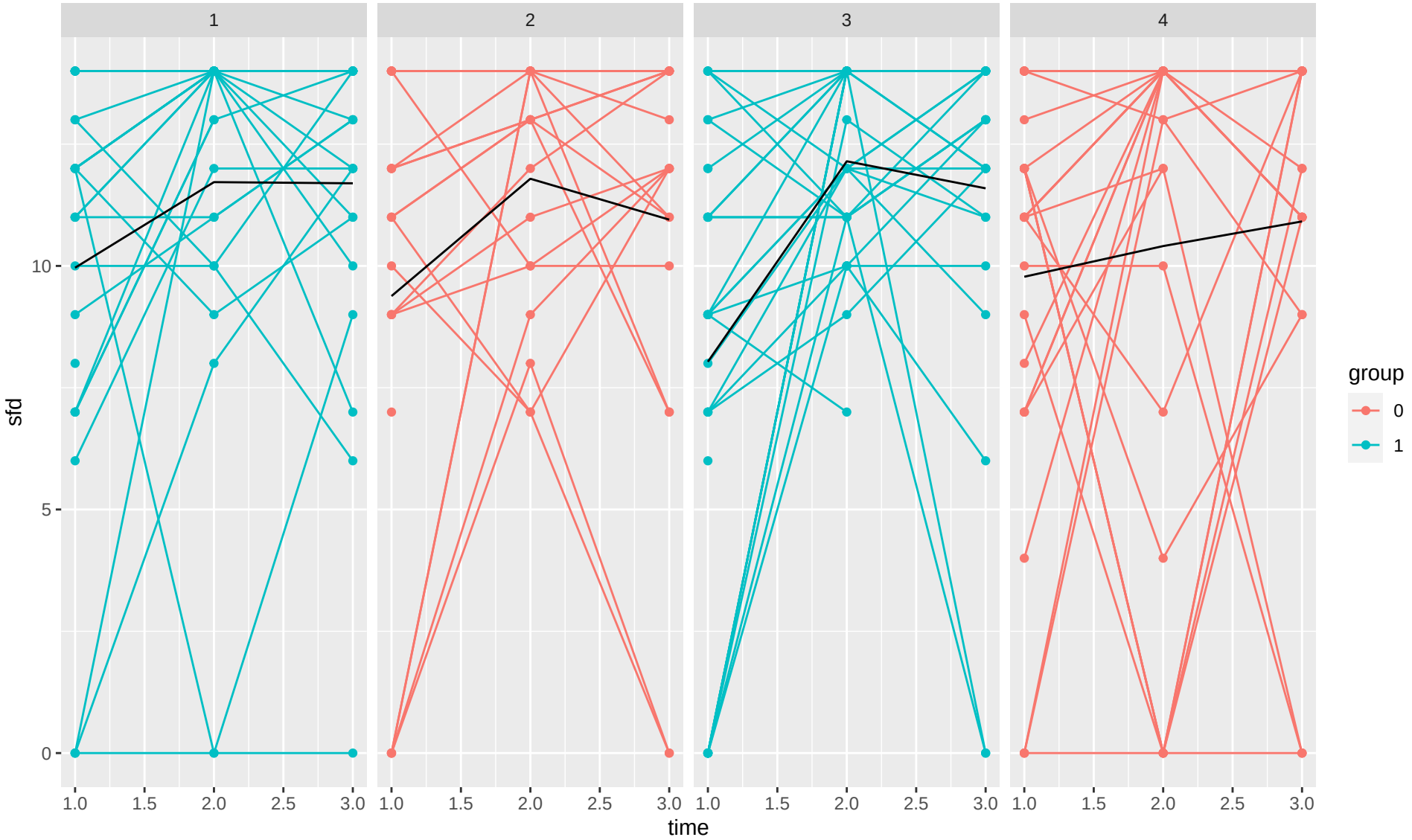
Variable	Definition
ID	Participant's ID
Time	Follow up time (Baseline, 6 months, 12 months)
Group	Intervention group (control or Intervention)
SFD	Symptom free days in the past two weeks
School	School recruited for the study

Table 1: Data Descriptions

Data Exploration



Data Exploration



Data Exploration

Current outcome: SFD (Count data)

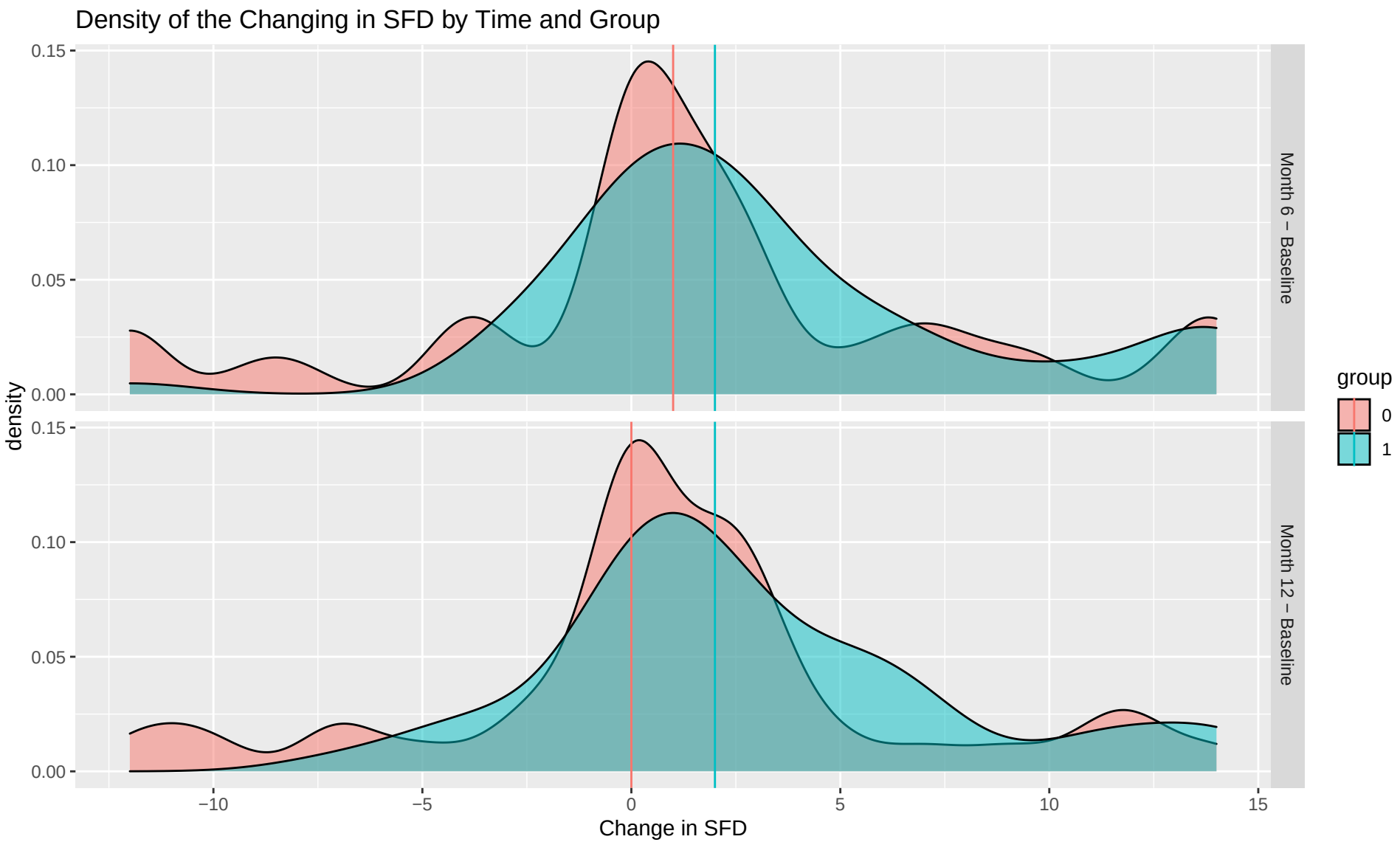
- Due to the skewed distribution towards higher values a poisson model will not fit our data well
- Outcome does not seem linear over the time observations.

Interested in the change from baseline to observation times.

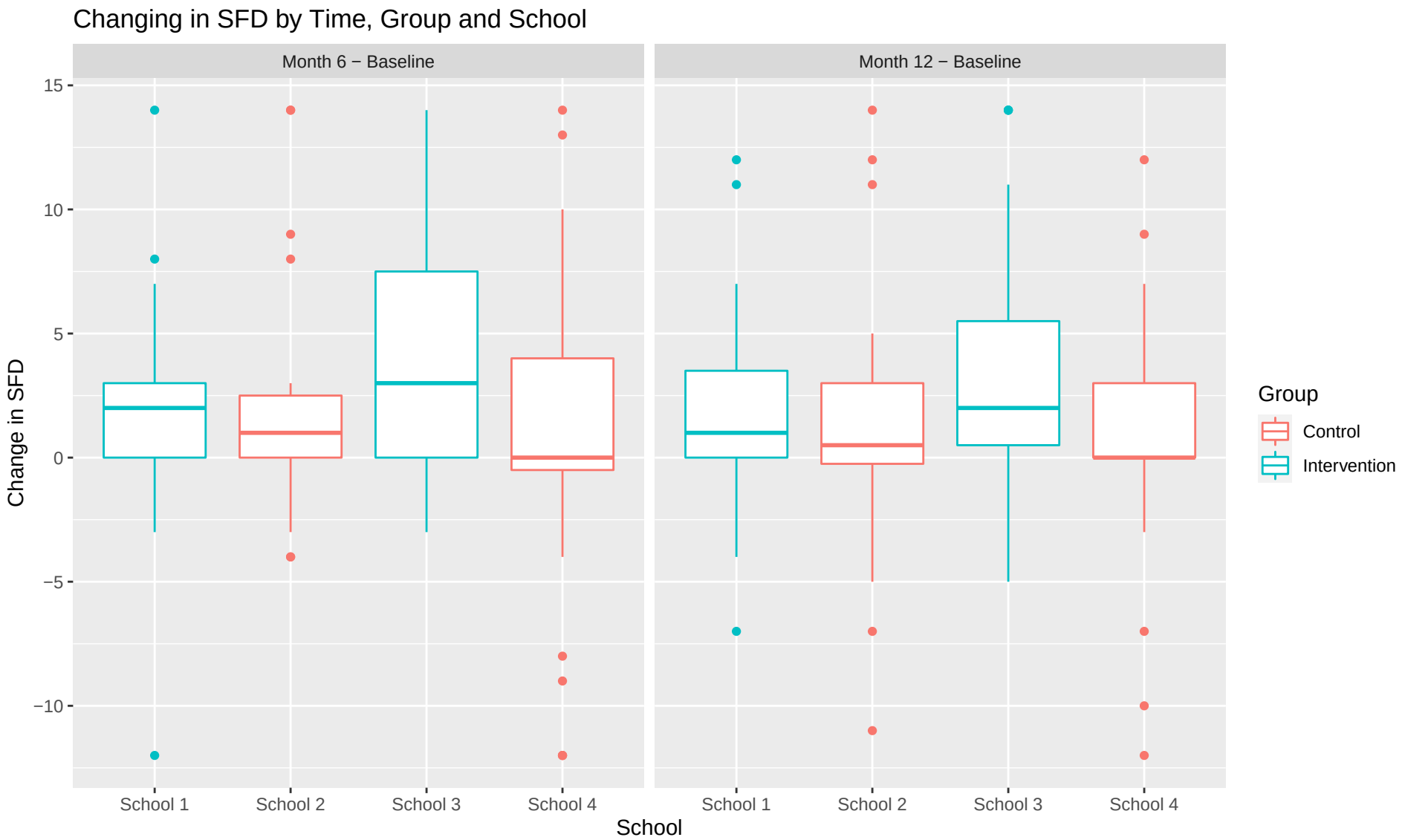
New outcome: Change in SFD (Continuous Data)

- Transform the SFD by calculated:
  - 6 month observation - baseline
  - 12 month observation - baseline
- Baseline with become covariate

New Continous Outcome



Exploring variation between school and within school

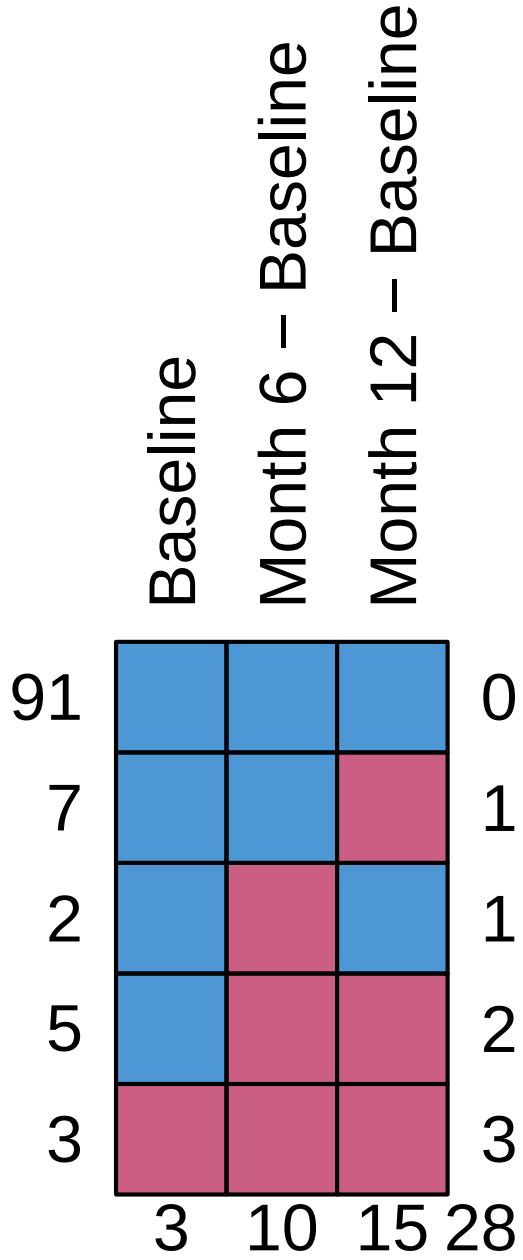


Data Description

Characteristic	Overall, N = 108[1]	Group Control		Group Intervention		p-value[2]
		School 1, N = 28[1]	School 3, N = 31[1]	School 2, N = 21[1]	School 4, N = 28[1]	
baseline	11.0 (7.0, 13.0)	12.0 (7.5, 13.0)	9.0 (6.2, 11.8)	11.0 (9.0, 14.0)	11.0 (7.5, 12.5)	0.3
Unknown	3	1	1	0	1	
Change_6months	2 (0, 5)	2 (0, 3)	3 (0, 8)	1 (0, 2)	0 (0, 4)	0.3
Unknown	10	3	4	2	1	
Change_12months	1.0 (0.0, 4.0)	1.0 (0.0, 3.5)	2.0 (0.5, 5.5)	0.5 (-0.2, 3.0)	0.0 (0.0, 3.0)	0.2
Unknown	15	5	4	1	5	

[1] Median (IQR)
[2] Kruskal-Wallis rank sum test

Missing Data



Model Specifications

To model change in SFD let i for school, j for subjects, k for measures.

We will used mixed effect model.

$$\begin{aligned} Y_{ijk} = \beta_0 &+ \beta_1 Baseline_{ij} \\ &+ \beta_2 Group_{ij} + \beta_3 Compare_{ijk} \\ &+ \beta_4 Group_{ij} \times Compare_{ijk} \\ &+ \alpha_{0i} + \alpha_{0j} + \epsilon_{ijk} \end{aligned}$$

where $\alpha_{0i} \sim N(0, \sigma_w^2)$, $\alpha_{0j} \sim N(0, \sigma_b^2)$, and $\epsilon_{ijk} \sim N(0, \sigma^2)$.

Missing Data Assumptions

We will be assuming data is missing at random (MAR)[1].

- MAR assumption: $R \perp\!\!\!\perp Y_{mis} | X, Y_{obs}$
- Separable parameter assumption
- Ignorability condition

$$
\begin{aligned}
L_i^{\mathrm{O}}(\theta, \psi) &\propto f_{\theta,\psi}\left(Y_{\mathrm{obs},i}, R_i, X_i\right) \\
&= f_{\psi}\left(R_i \mid Y_{\mathrm{obs},i}, X_i\right) f_{\theta}\left(Y_{\mathrm{obs},i} \mid X_i\right)
\end{aligned}
$$

Model Result

Fixed Effects Estimates:

Characteristic	Beta	95% CI[1]	p-value
baseline	-0.82	-1.0, -0.69	<0.001
group			
0	—	—	
1	1.1	-0.46, 2.6	0.2
compare			
m6_m0	—	—	
m12_m0	-0.05	-1.4, 1.3	>0.9
group * compare			
1 * m12_m0	-0.27	-2.2, 1.6	0.8

[1] CI = Confidence Interval

Random Effects Estimates:

group	Std.Dev	Variance
id	2.055	4.221
school	0.000	0.000
Residual	3.250	10.560

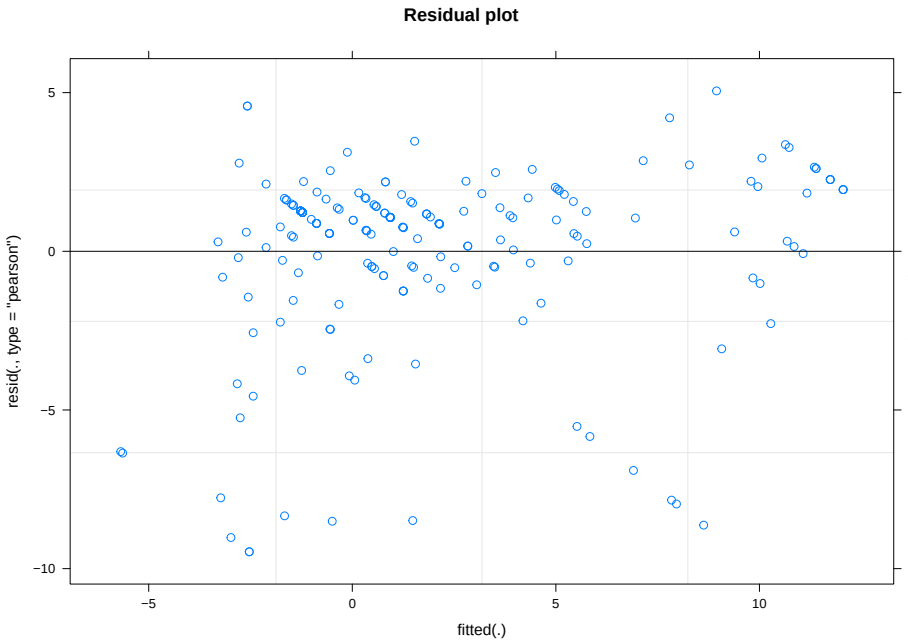
Model Interpretations:

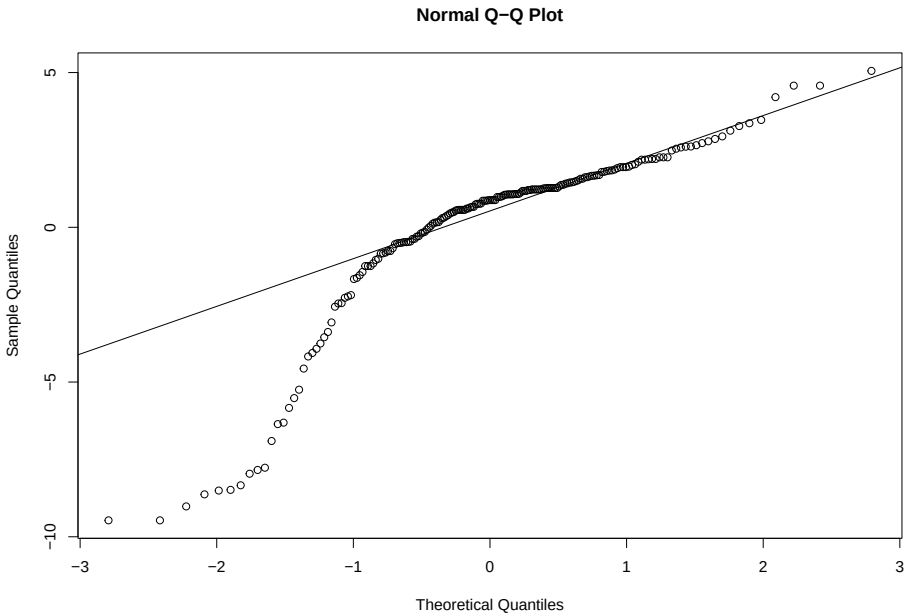
1. interpretations
2. interpretations
3. interpretations

Note: This model is singular

- Conclusion:
  - When comparing the 6 months and baseline, The increase of sfd in intervention group is 1.1(-0.46, 2.6) more than the increase of sfd in

Model Quality





Cluster RCT design

- The investigators wish to propose a **cluster-randomized clinical trial (RCT)** in 30 Bronx schools to evaluate the effectiveness of their intervention program.
- Primary hypothesis: compared to the control group, children in schools randomized to intervention group will experience a greater improvement in the number of SFD at **any of the 3, 6, 9, and 12 months** assessment.
- The investigators would like to have at **80% probability** to declare the trial is successful if the true effect size in **improvement of SFD over time** is **at least 1/3 standard deviation**.

Study design proposal:

- The primary interest of the study: test whether there is difference at any of the 3,6,9,12 months;
- Consider 4 comparisons separately:
  - month3: sfd_chage~group
  - month6: sfd_chage~group
  - month9: sfd_chage~group
  - month12: sfd_chage~group

Two levels of Sampling

- N_1: Number of Indiviuals in each School *(What we want to estimate)*
- N_2: Number of Schools for one treatment arm (15 in our case)

Study design proposal:

2 level structure[2]:

$$y_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 y_{0,ij} + u_i + \epsilon_{ij}$$

- i for school, j for subject, y_{ij} is the sfd_change from tested time point to baseline
- within each comparison, there is only 1 measurement. So no consideration of the intra-subject correlation
- $u_i \sim N(0, \sigma_2^2)$, random intercept between schools
- $\epsilon_{ij} \sim N(0, \sigma_e^2)$, random error term

Intraclass Correlation Coefficient

$$\rho_1 = \frac{\sigma_2^2}{\sigma_2^2 + \sigma_e^2}$$

Hypothesis Set Up

Hypothesis: $H_0 : \beta_1 = 0, H_1 : \beta_1 \neq 0$

- calculate N based on normal distribution, with multiple adjustment: $\alpha^* = \alpha/4 = 0.025/4$
- $\beta = 0.2$
- $N_1 = N_0 = 15$
- Interested in when stadardized effect size $\Delta = 1/3$

Sample Size Calculation

Test statistics

$$D_2 = \frac{\sqrt{N_2 N_1}\left(\bar{Y}_1 - \bar{Y}_0\right)}{\sigma\sqrt{2(1-\rho)}} \sim N(0,1)$$

Sample Size formula

$$N_1 = \frac{2(1-\rho)z^2_{\alpha^*,\phi}}{N_2\Delta^2_{(2)} - 2\rho z^2_{\alpha^*,\phi}}$$

Where z is calculated based on the normal distribution.

$$z^2_{\alpha^*,\beta} = (z_{\alpha^*/2} + z_\beta)^2 = \left[\Phi^{-1}(1-\alpha^*/2) + \Phi^{-1}(1-\beta)^{-1}\right]^2$$

Intra Class Correlation in our Models

Model 1 Random Effects:

group	Std.Dev	Variance
id	2.055	4.221
school	0.000	0.000
Residual	3.250	10.560

1. $\rho_1 = 0.286$
2. $\rho_2 = 0.000$

Model 2 Random Effects:

group	Std.Dev	Variance
id	4.384	19.215
school	0.597	0.357
Residual	3.275	10.726

1. $\rho_1 = 0.646$
2. $\rho_2 = 0.012$

Sample Size Suggested

rho	class_size	group_size	total_samp
0.00	13.381	200.720	401.440
0.01	15.294	229.411	458.822
0.03	21.685	325.278	650.556

Resources

[1] Hogan, J. W., Roy, J., & Korkontzelou, C. (2004). Handling drop-out in longitudinal studies. Statistics in Medicine, 23(9), 1455–1497. https://doi.org/10.1002/sim.1728

[2] Ahn, C., Heo, M., & Zhang, S. (2014). Sample size calculations for clustered and longitudinal outcomes in clinical research. CRC Press.