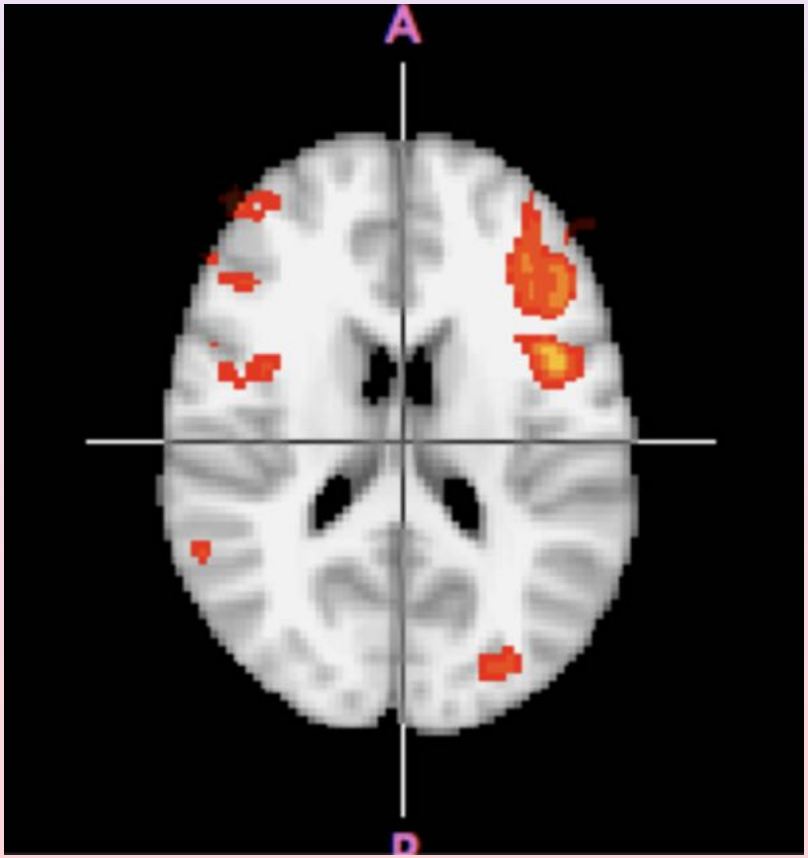
Evidence-based Second-Generation p-values on Functional Magnetic Resonance Imaging Data

Ya-Chen Lin

Department of Biostatistics
Vanderbilt University
March 23, 2020

Introduction to fMRI data

1. Task-induced fMRI
   - Participants are engaged with tasks
   - Look for localized brain activation patterns
2. Statistical inference relies on p-values



P-values as an inference tool - good or bad?

Advantages:

- Simple computation
- Widely used in all fields

Disadvantages:

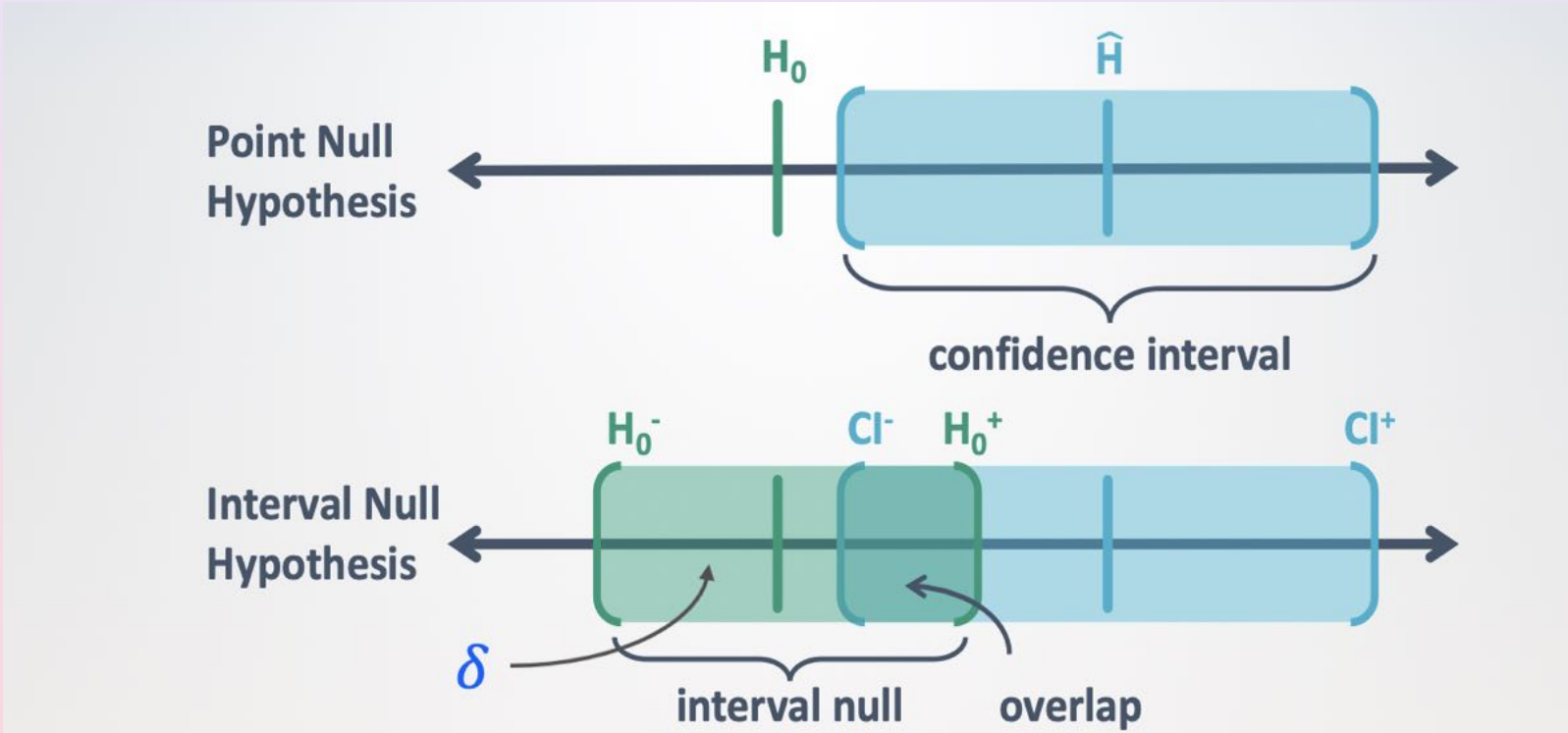
- Interpretation issues[1]
- Statistical v.s. Clinical significance[2]

---

[1] Hubbard et al. (2003)

[2] Mark et al. (2016); Ranganathan et al. (2015)

SGPV brings in interval null testing

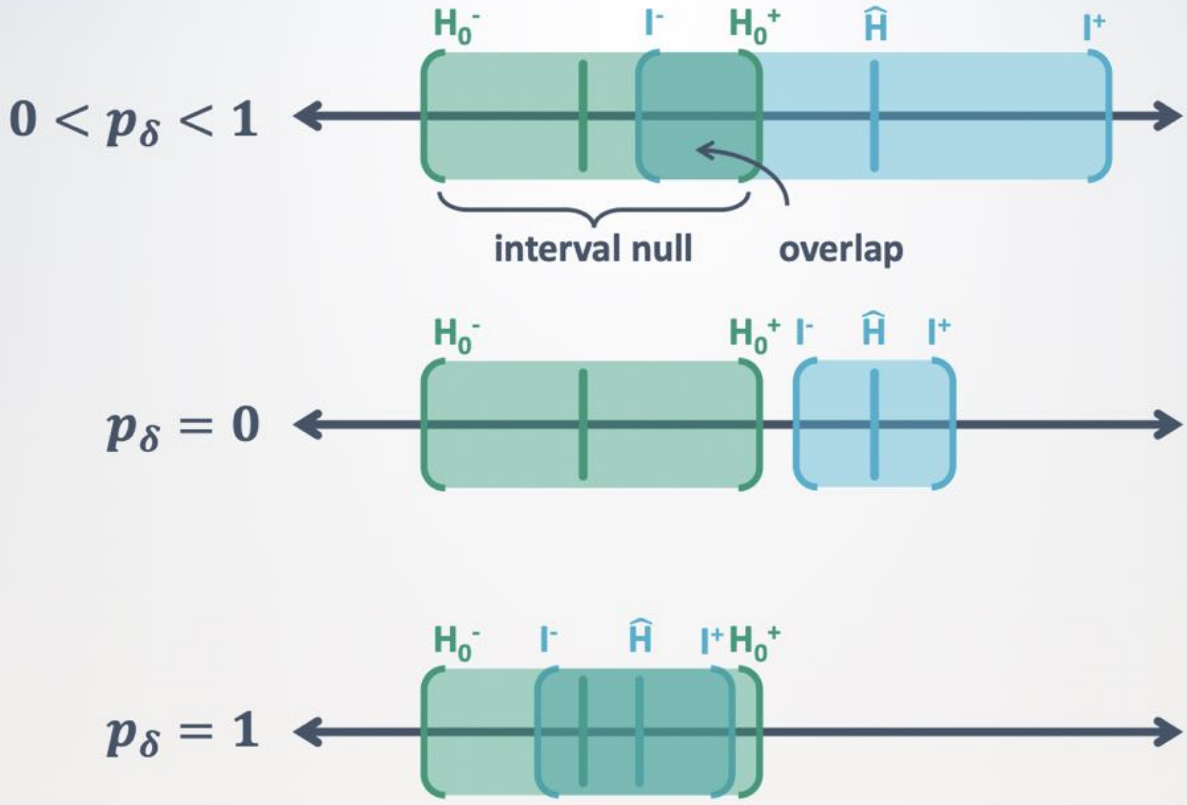


Point null hypothesis H_0 and interval null hypothesis $[H_0^-, H_0^+]$

Data-supported hypothesis $\hat{H}$ and confidence interval $[CI^-, CI^+]$

[1] Dr. Jeffrey Blume's slides

Simple interpretation as fraction of overlap



Works with confidence, credible, and support intervals

[1] Dr. Jeffrey Blume's slides

The computation of SGPV is concise

Second-generation *p*-value (SGPV)

$$p_\delta = \frac{|I \cap H_0|}{|I|} \times \max\left\{\frac{|I|}{2|H_0|}, 1\right\}$$

Proportion of data-supported hypotheses that are also null hypotheses

Small-sample correction factor

shrinks proportion to ½ when $|I|$ wide

when $|I| > 2|H_0|$

[1] Dr. Hakmook Kang's slides

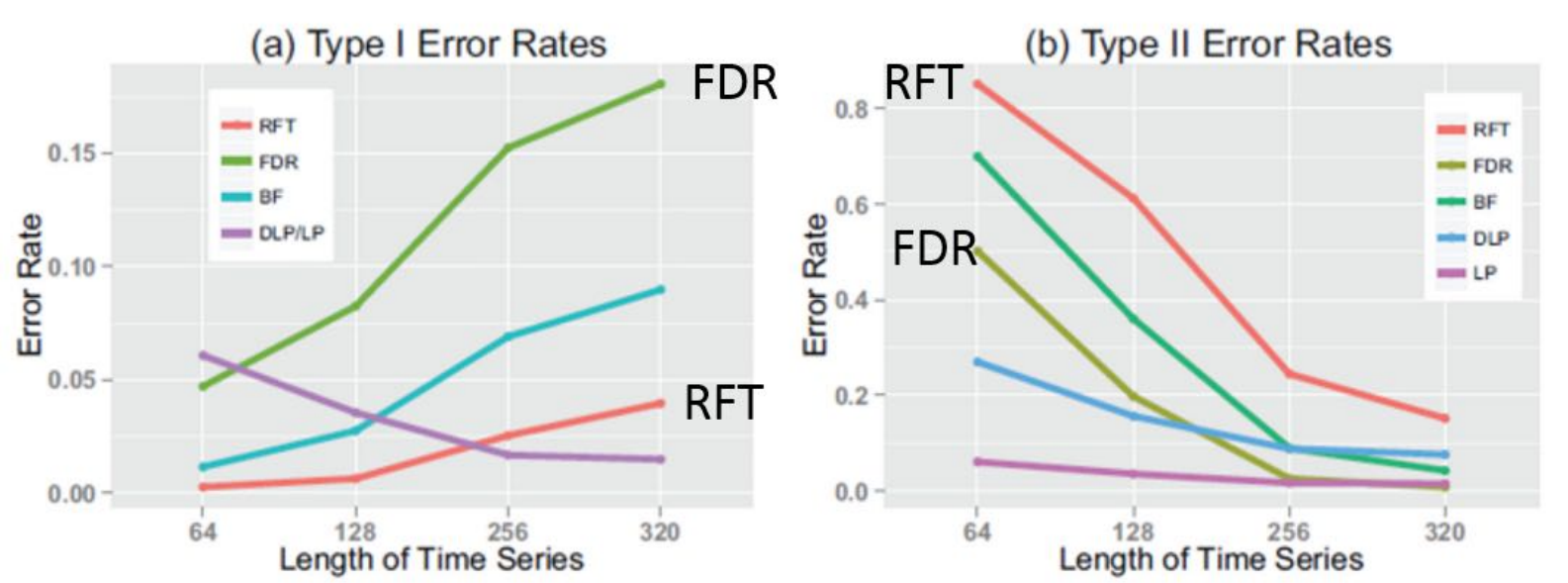
Balancing power and Type I error rate is challenging in fMRI analysis

Nature of the data

- Conduct analysis on large number of voxels
- Data are noisy

Multiple comparison adjustments

- Control FWER (Random Field Theory (RFT))
- Control False Discovery Rate (FDR)

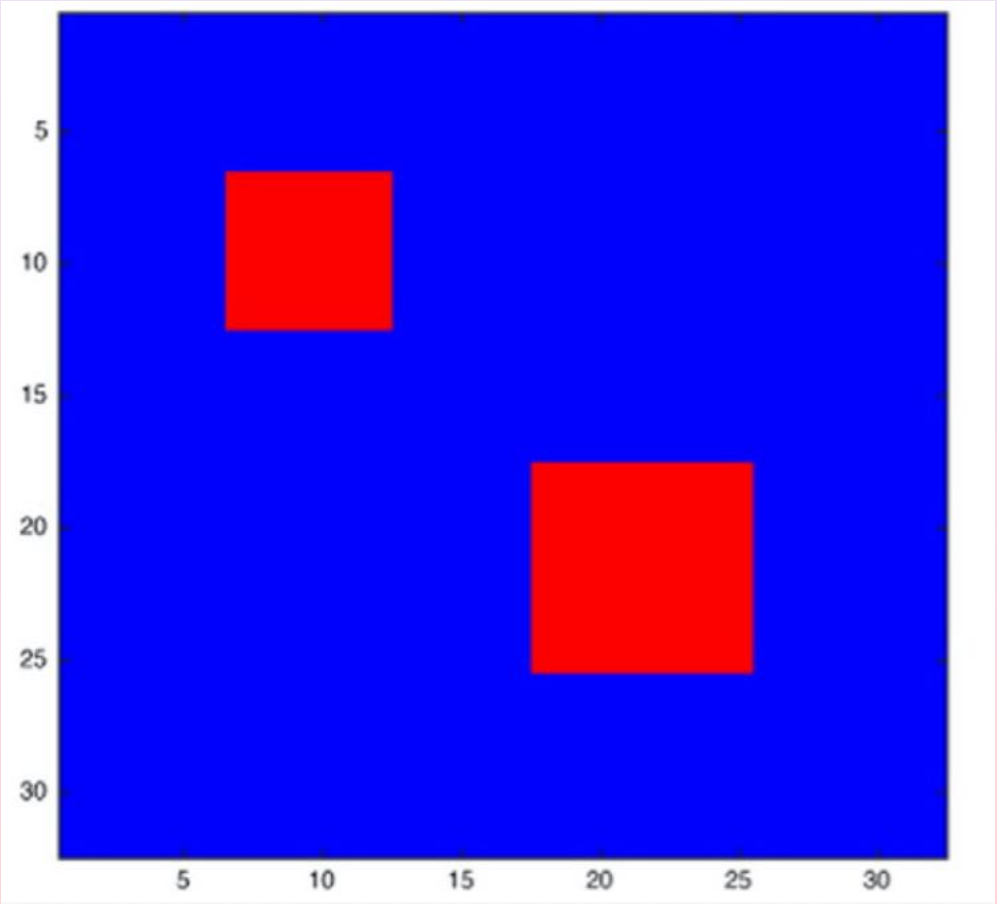


[1]Single subject simulation results from Kang et al. (2015)

Settings to simulate task-induced fMRI data

- Spatio-temporally correlated data
- 32×32, 2-D images with two active regions
- At each voxel and a time, with P stimuli,
  $Y_v(t) = \Sigma_{p=1}^{P} X_p(t)\beta_v^p + \epsilon_v(t)$
- Clinically null region
  - No neurons in Cerebrospinal fluid (CSF)
    $\rightarrow$ any signal is noise
  - 10×10 region
  - $\beta^1 = \beta^2 = 0$
- Vary time lengths and sample sizes
- Gaussian kernel with FWHM = 8 mm

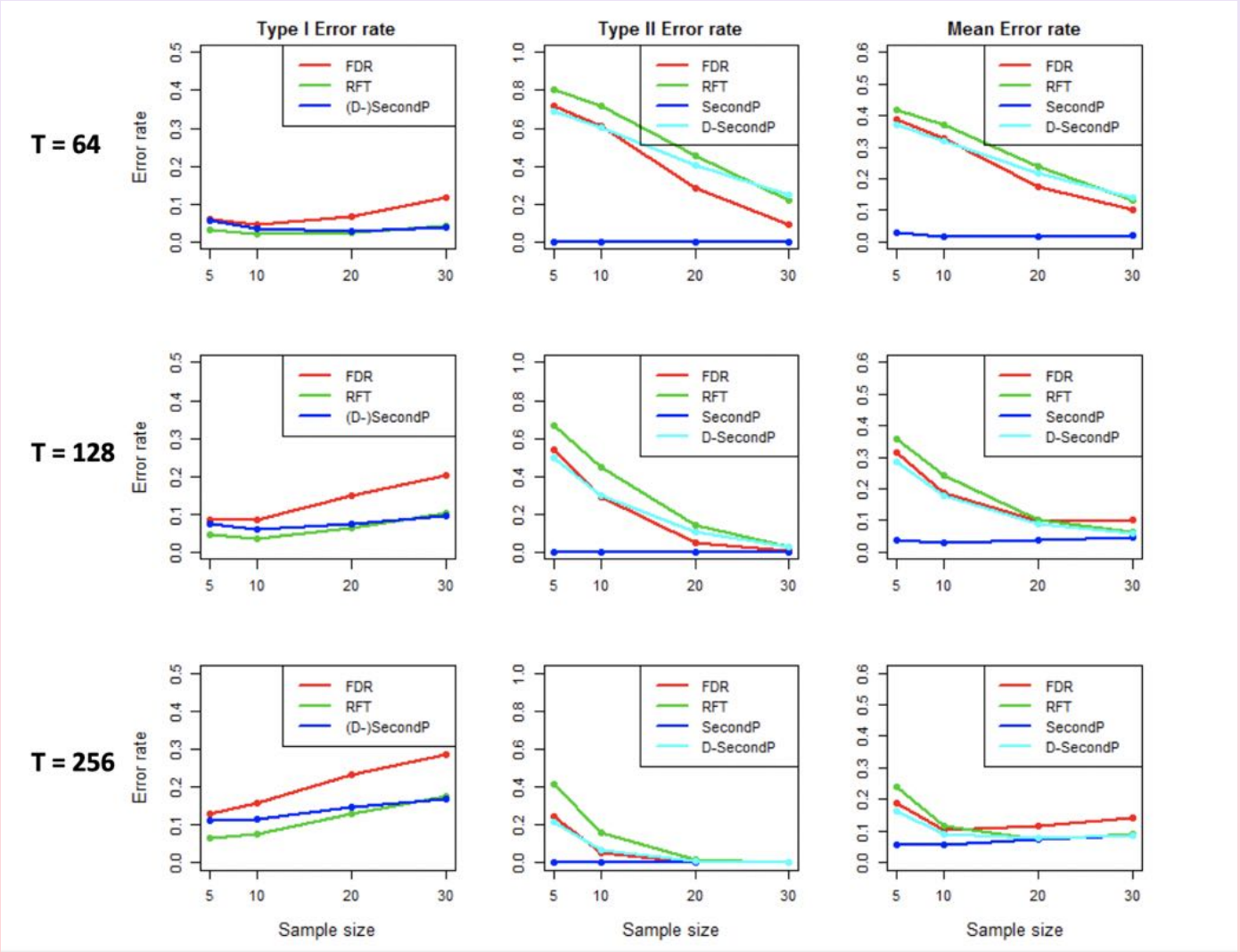
Visualize the true image



Estimation and Inference of simulation

1. Fit the linear model at each voxel to estimate β^2 - β^1
2. SGPV clinical null region:
   - $0 \pm$ IQR(all $\hat{\beta}^2 - \hat{\beta}^1$ in CSF) / 6
3. Methods compared:
   - SGPV: Compute SGPV with 95% CI
   - FDR : Compute p-values; control at FDR = 0.05
   - RFT : Compute z-scores; control error probability at 0.05
4. Fair comparison: Dichotimized SGPV (D-SecondP)
   - Voxels with SGPV = 0 are deemed as significant
   - Voxels with SGPV = 1 are deemed as insignificant
   - Voxels with $0 <$ SGPV < 1 (inconclusive region) are deemed as insignificant

SGPV obtains power and keep Type I error rate steady



Participants and task in real fMRI data

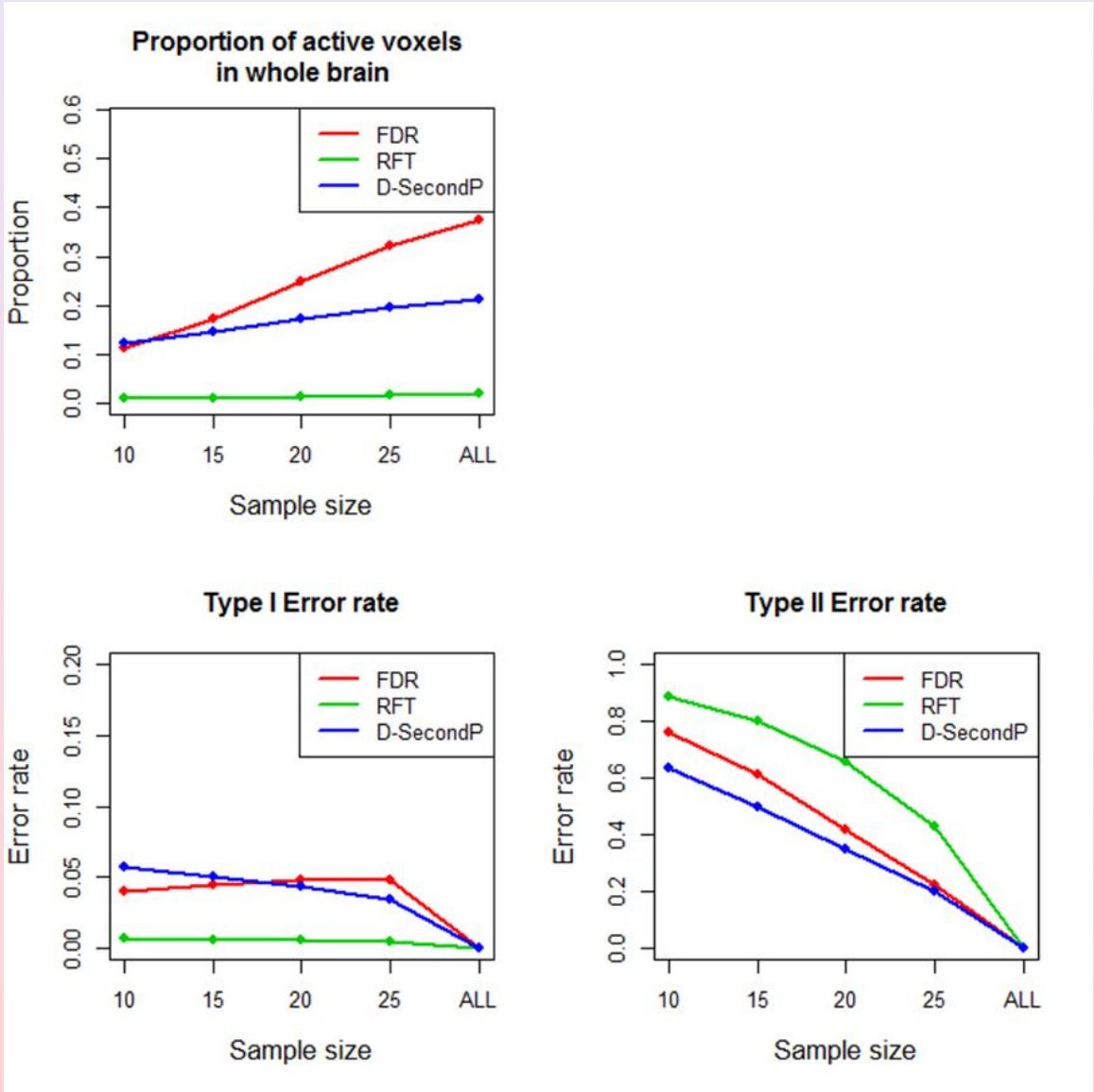
Data:

- 29 women with major depression disorders, aged between 45 and 75
- Spatial attention task that measures attention bias

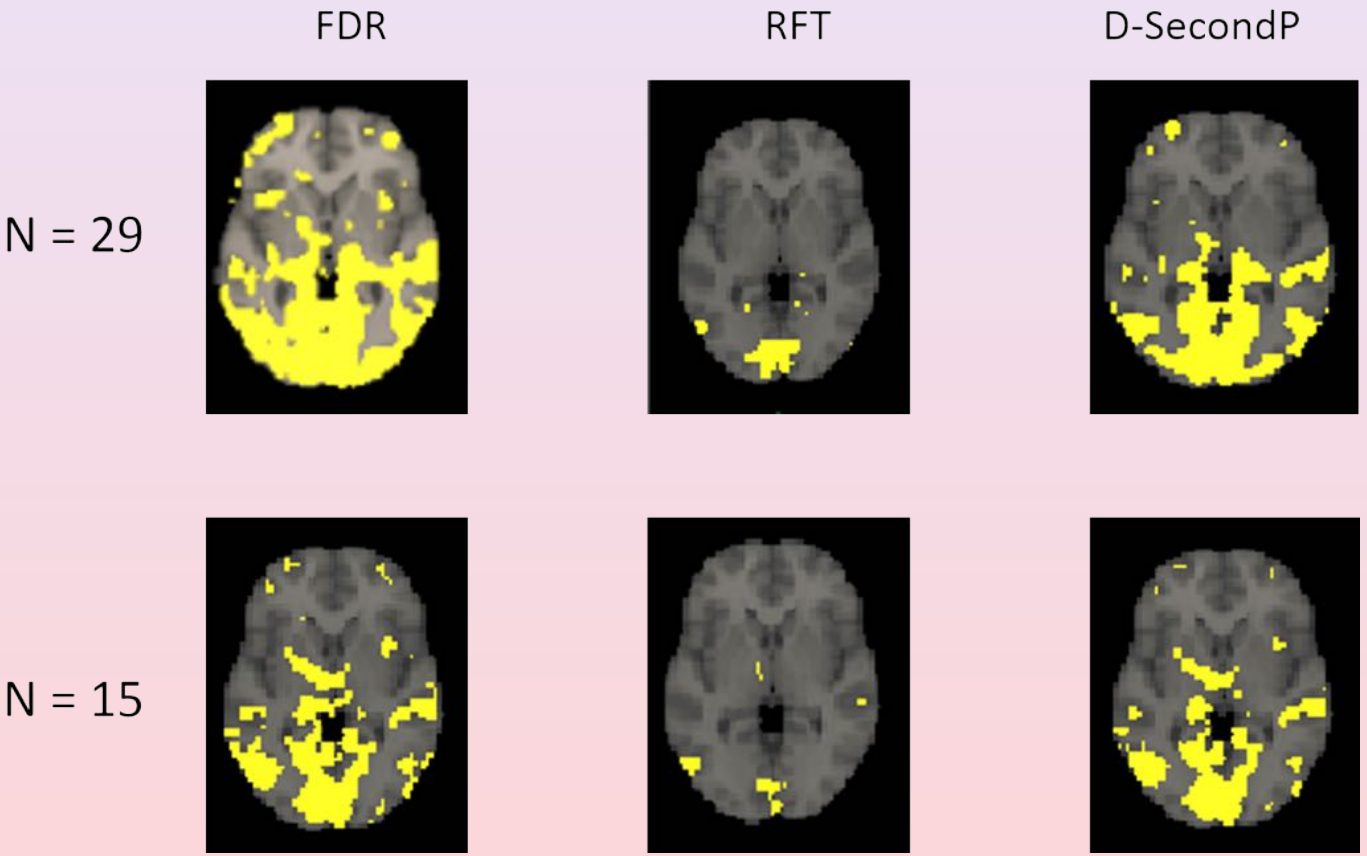
Analysis:

1. Linear model at each voxel to estimate $\beta^2 - \beta^1$
2. Compute SGPV with clinical null region
   - $0 \pm$ IQR/6 using estimates from voxels in CSF
3. For voxels outside of CSF
   - Compute SGPV with 95% CI
   - Control z-scores with RFT at 0.05
   - Control p-values with FDR at 0.05
4. Data Decimation

SGPV method is the most robust method



Visualization of the results



SGPV offers good properties in fMRI analysis

1. Convenient and simple interpretation
2. Incorporate clinical information (in fMRI: data collected from CSF)
3. Provide transparency, rigor and reproducibility of scientific results
4. Reduce false positives in fMRI data analysis

Tools have been developed; more are coming

1. Method:
   - "*An Introduction to Second-Generation p-Values*", Blume et al. (2019)
   - "*Second-generation p-values: Improved rigor, reproducibility, transparency in statistical analyses*", Blume et al. (2018)
2. R package:
   - github.com/weltybiostat/sgpv
3. R-shiny app for fMRI analysis (work in progress)
   - Offer clear explanation of SGPV
   - Provide interactive visualization
   - Download results with format compatible with other imaging software

Acknowledgments

Advisor: Prof. Hakmook Kang, Department of Biostatistics, Vanderbilt University

SGPV method:

- Prof. Jeffrey Blume, Department of Biostatistics, Vanderbilt University
- Valerie Welty, Department of Biostatistics, Vanderbilt University

Collaborators:

- Dr. Warren Taylor, Department of Psychiatry, Vanderbilt University
- Brian Boyd, Department of Psychiatry, Vanderbilt University