Multiple Comparisons:
Embracing Instead of Fighting!

Gang Chen

Scientific and Statistical Computing Core

National Institute of Mental Health

National Institutes of Health, USA

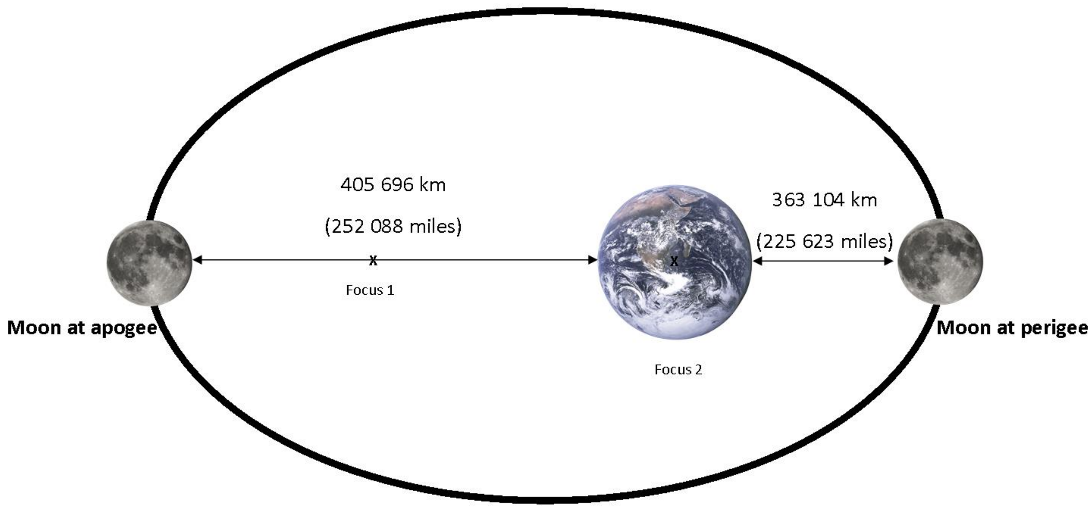
Handout: afni26_ROI-based-modeling.pdf

Preview

- **Current correction methods for multiplicity**
- **3 perspectives**
  - NHST: *p*-value and thresholding
  - Model accuracy
  - Integrative modeling
  - 2 toy examples: NBA players; Kidney cancer
- **Application: region-based analysis (RBA)**
  - Program in AFNI: ***RBA***
- **Other applications**
  - Matrix-based analysis (program in AFNI: ***MBA***)
  - Region-based inter-subject correlation (ISC) analysis
  - Gray matter connectivity analysis (DTI)
  - Other cases involving multiplicity

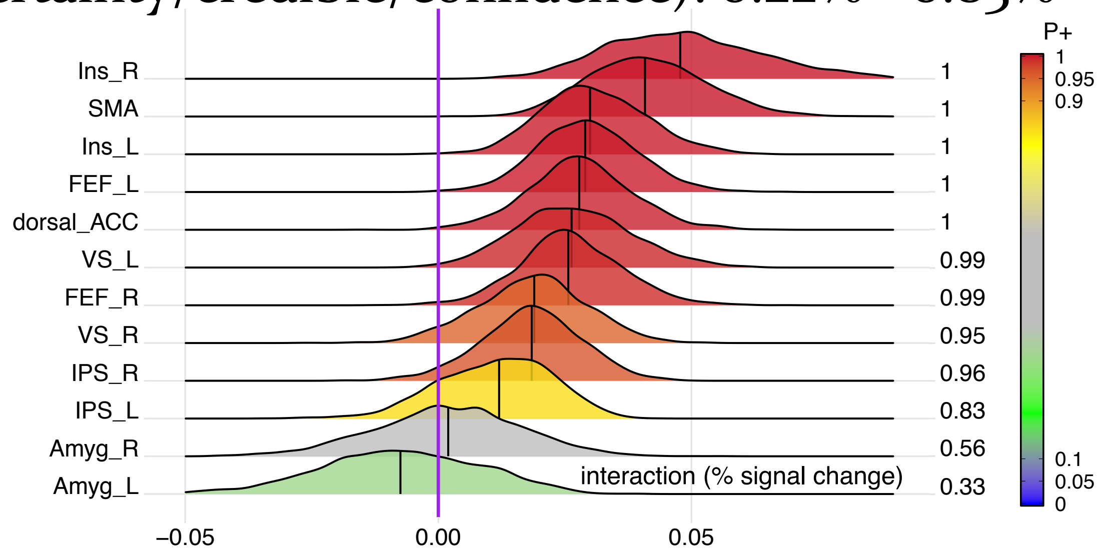
Reproducibility: start with physics

- **What is the distance between earth and moon?**
  - *t*-statistic = 4.25 (or *p*-value = 0.01): informative?
  - Ridiculous? Check out colorbars, tables and network graphs in publications/slides/posters…
  - Average: 384,400 km
  - Uncertainty:
    363,104 km – 405,696 km



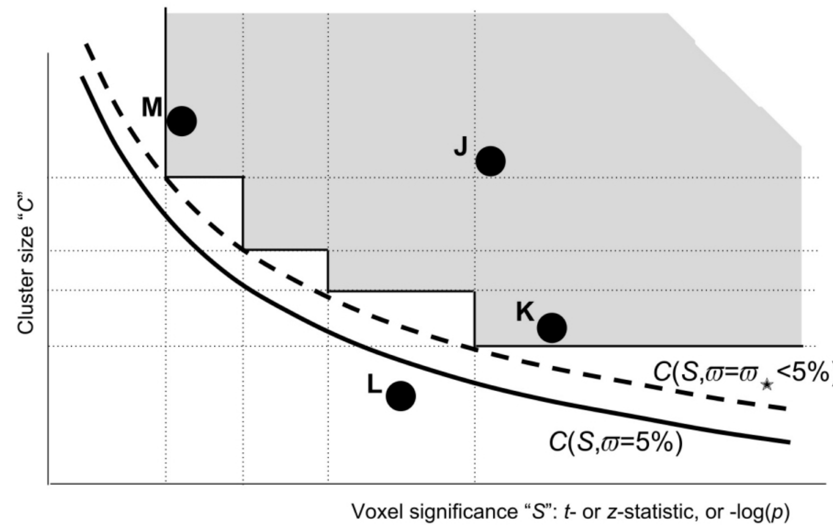
Reproducibility: neuroimaging

- **What is the BOLD response in a brain region?**
  - *t*-statistic = 4.25 (or *p*-value = 0.01): informative?
  - Colorbars, tables and network graphs in publications/slides/posters…
  - Average: 0.52% signal change
  - Range (uncertainty/credible/confidence): 0.22% - 0.83%



Multiplicity: correctness in the eye of beholder

- **100,000 spatial units - 100,000 models: MUA**
  - Assumption of spatial independence
  - Sharing no information
- **Corrections**
  - Multiplicity + spatial relatedness
  - Heavy penalty: information waste
  - Arbitrariness
    - Why not 0.04 or 0.06 instead of 0.05?
    - Different correction methods: arbitrary voxel *p* vs. power
    - Heavily dependent on data space: whole brain, gray matter, ROIs
    - Information waste at global level: only local relatedness considered



Research reproducibility

- **Does strength of statistical evidence shrink?**
  - Previous claim with statistical evidence: p-value = 0.03
  - Current study with evidence: p-value = 0.04.
  - Multiple testing issue? Should one adjust for multiplicity?
  - How about all studies that use statistical analyses?

- **How are study repetitions distributed?**
  - Same experiments repeated 100 times
  - An effect (population, BOLD) across entities (counties, brain regions)

Null Hypothesis Significance Testing

- **Straw man H_0: null hypothesis**
  - Witch hunt: Don Quixote's windmills
  - <u>**Type I error**</u> = P(data | H_0) = false positive = *p*-value
    - Surprise or weirdness of data: **0.05**
    - No effect until shown with small *p*-value
    - Innocent until proven guilty
  - <u>**Type II error**</u> = P(accept H_0 | H_1) = false negative

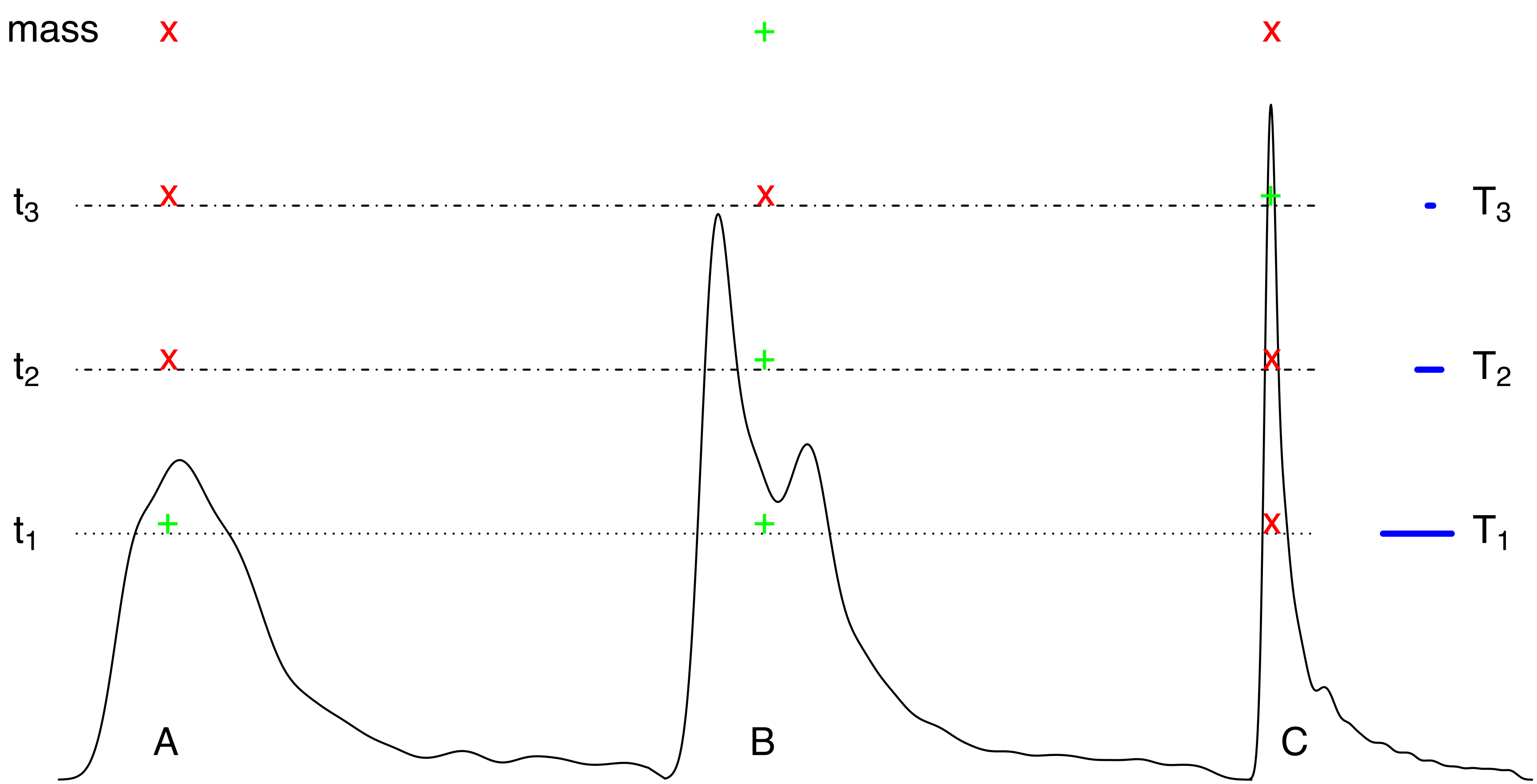
- **Real practice: type I error ONLY**
  - False positives: purely pleasing to statisticians!
  - With NO regard for type II error

	H_0 True	H_0 False
Reject H_0	Type I Error (false positive)	Correct
Fail to Reject H_0	Correct	Type II Error (false negative)

Results interpretation

- What is the conclusion of a region where p=0.6?
- If p=0.05, what is the probability for the region being activated?

Clusters vs islands: arbitrariness

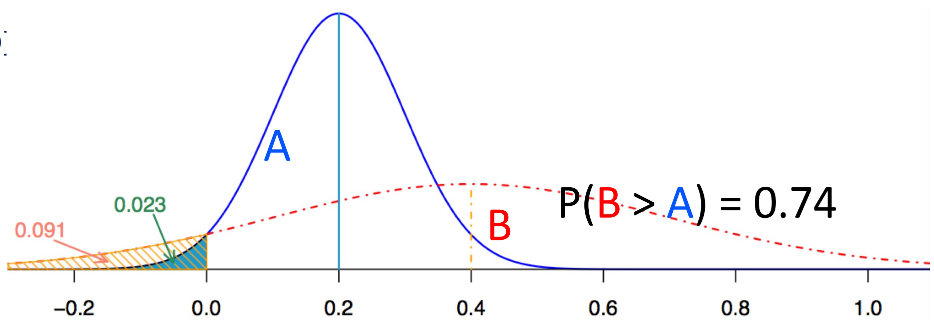


Issues: NHST

- **Arbitrary dichotomy:** **where to draw a line in the sand?**
  - Binary or discrete: innocent vs guilty
  - *p*-value itself is a random variable
  - Unrealistic: “activated” vs “not activated”?
  - Methods for correlation matrix: why is 0.3 so special?
- **Vulnerable to misconceptions**
  - **p (weirdness | H_0) ≠ p (H_0 | data)**
  - **Absence of evidence ≠ evidence of absence**
- **Vulnerable to data manipulations**
  - **Statistical evidence changes: whole brain, gray matter, region**
- **Inflated effect estimates**
  - **Type M (magnitude) and type S (sign) errors: biasedness**

Issues: NHST

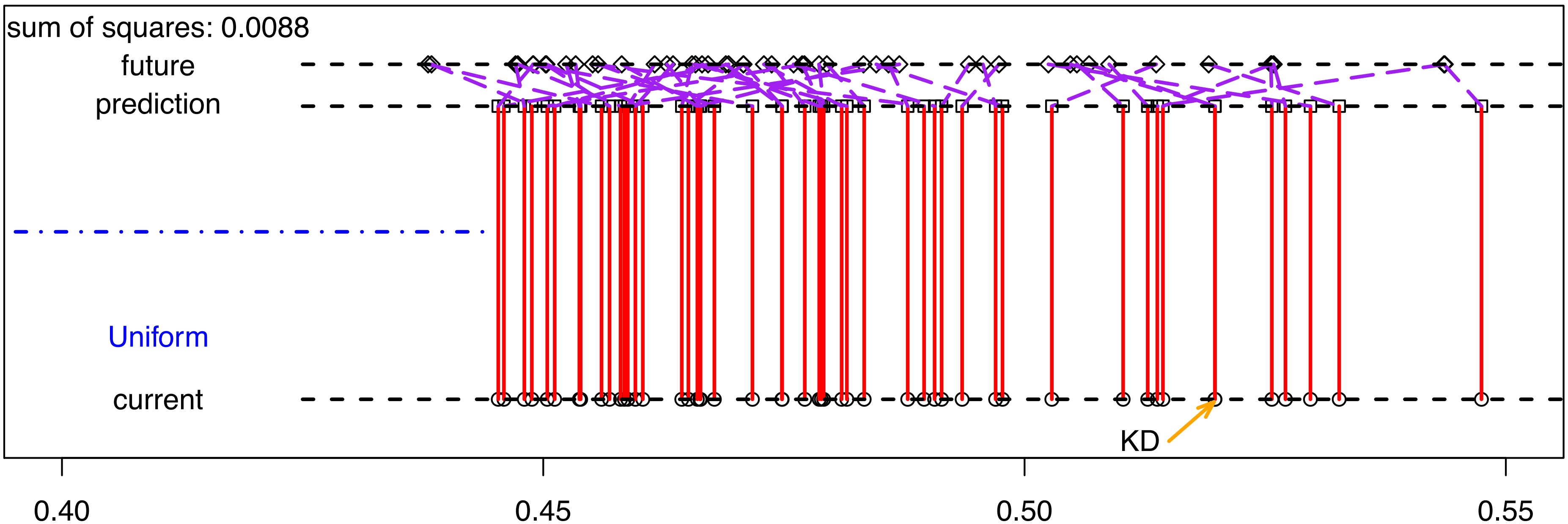
- **Inefficient modeling**
  - Over-penalizing
  - Ignore false negative (power)
  - No mechanism to incorporate prior knowledge
- **Disregarding effect size**
- **Uncertainty unavailable**
  - No standard deviation at voxel o
- **Lack of spatial specificity**
  - Locating regions per peak voxel
- **Penalizing small regions**



Chen, et al, 2019. Fighting or Embracing Multiplicity in Neuroimaging? Neighborhood Leverage versus Global Calibration. NeuroImage (in press)

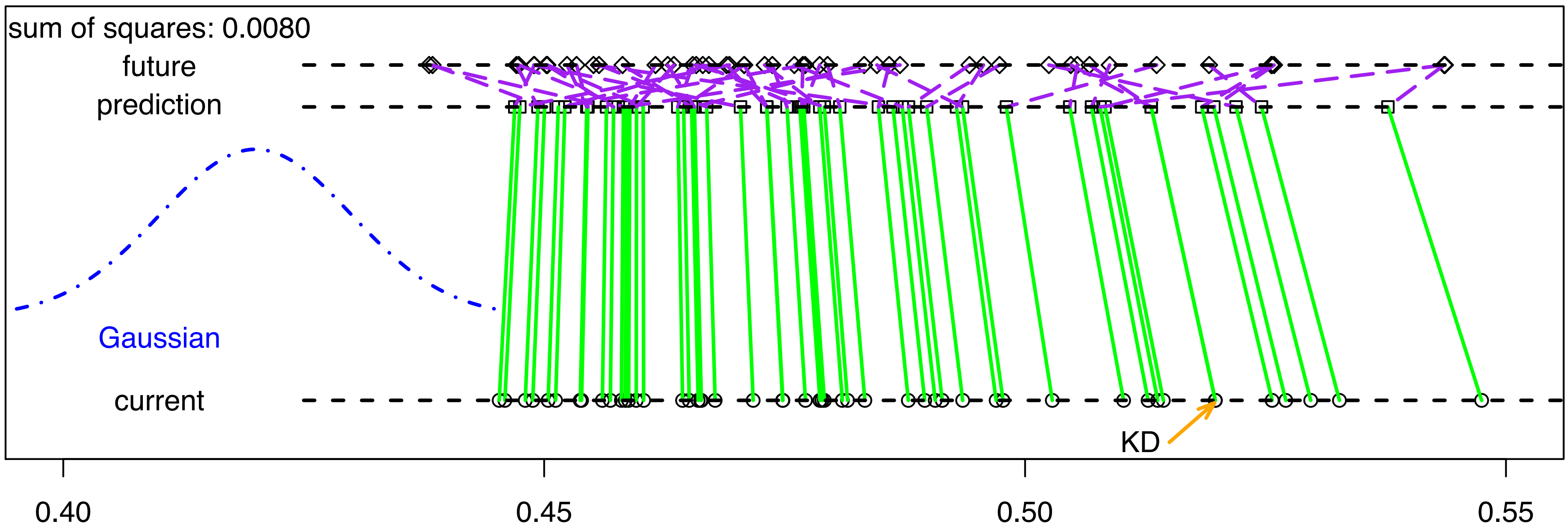
Toy Example 1

- **NBA players**
  - Kevin Durant field goals percentage: 52.1%
  - Prediction: performance during next season?
  - One vs. top 50 players: **no pooling** vs complete pooling



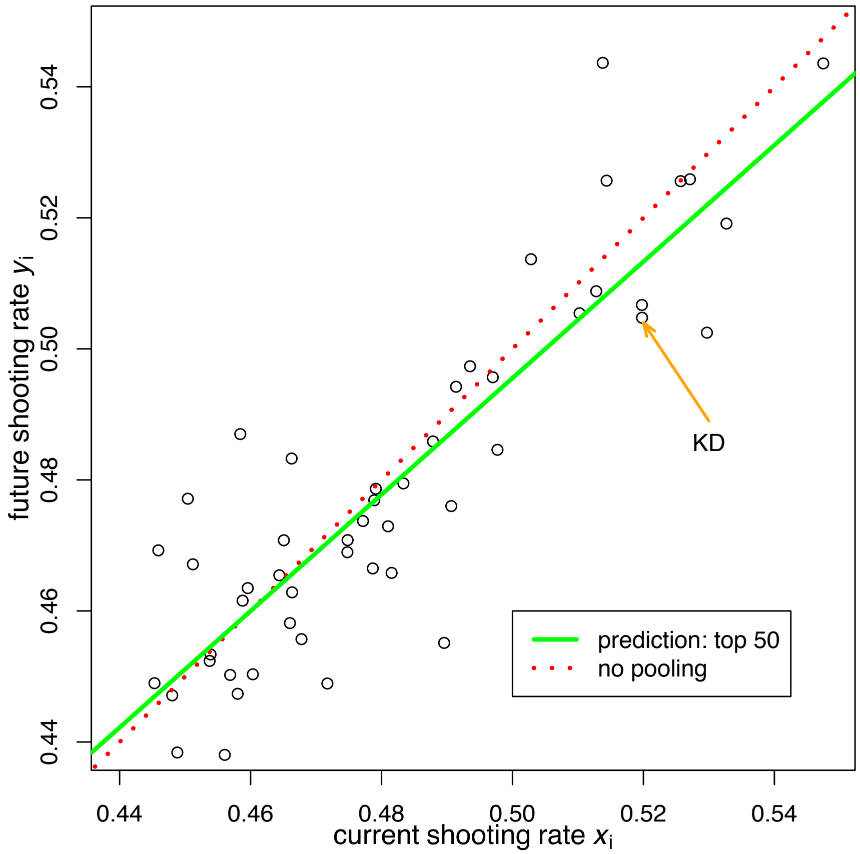
Toy Example 1

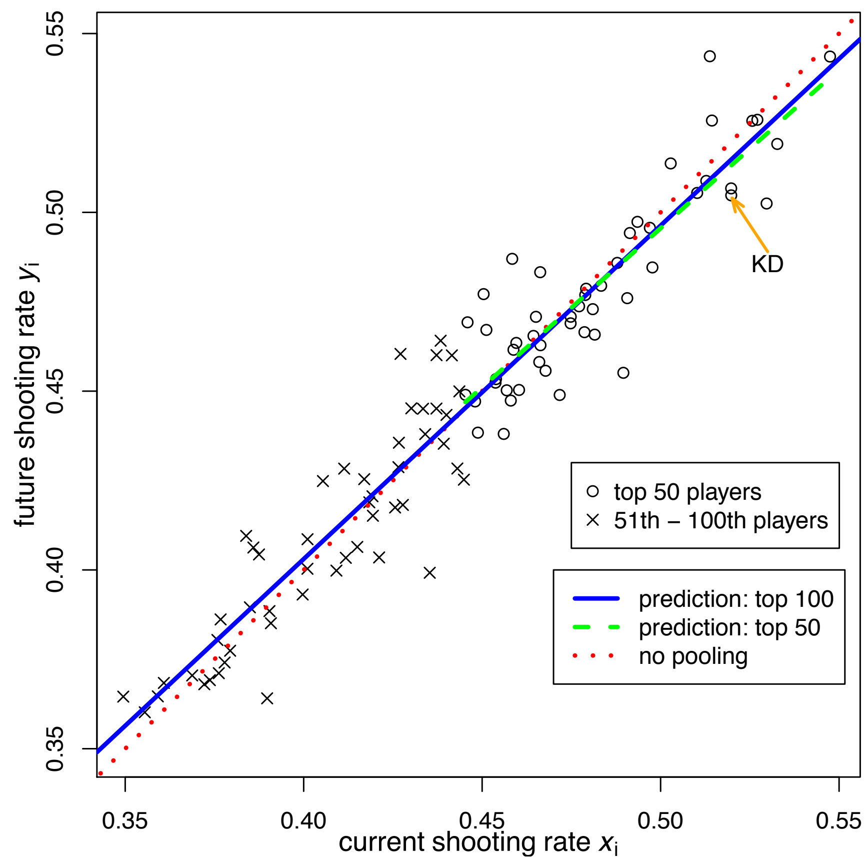
- **NBA players**
  - Kevin Durant field goals percentage during 2019: 52.1%
  - Prediction: performance during 2020?
  - One vs. top 50 players: **partial pooling** (regression to the mean)



Toy Example 1

- **Top 50 vs. 100 NBA players: adaptivity**

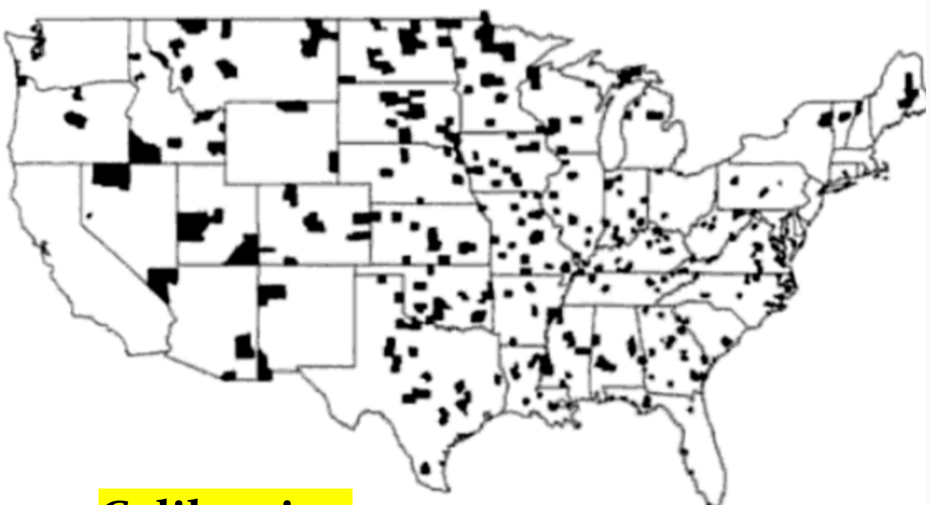




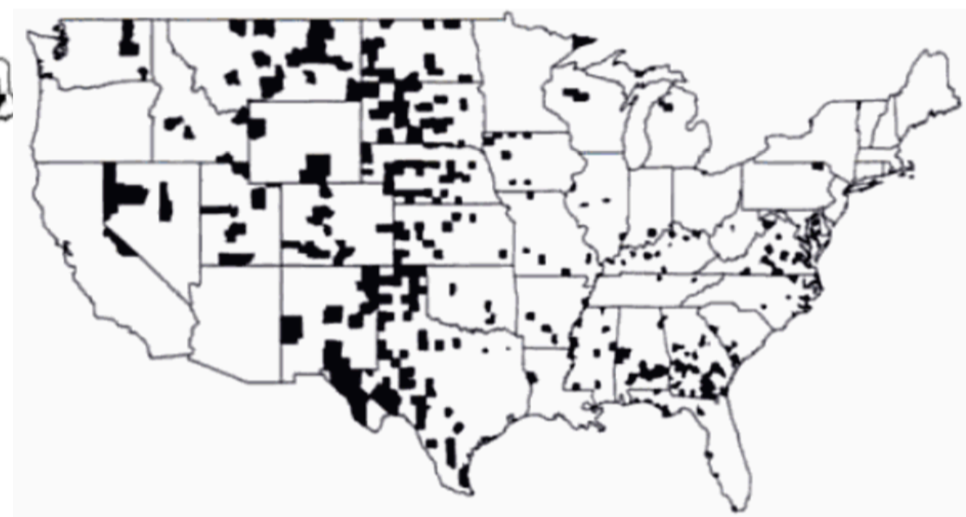
Toy Example 2

- **Kidney cancer distribution among U. S. counties**

Highest rate



lowest rate



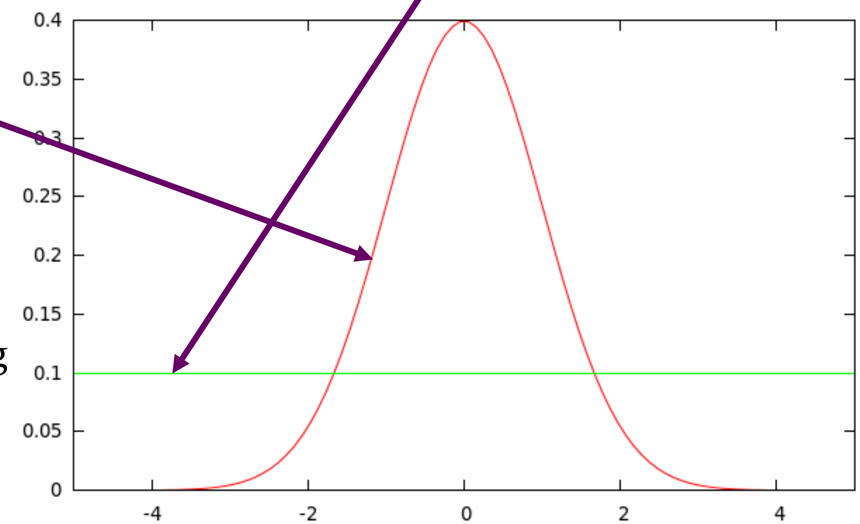
Calibration

Morals from kidney cancer data

- **Multiplicity problem: > 3000 counties!**
  - Divide *p*-value by number of counties?
  - Borrow idea from neuroimaging: leverage geographical relatedness?
- **What can we learn from the example? Food for thought**
  - Care about strawman H_0 (zero kidney rate), false positives, *p*-value?
  - Trust individual county-wise estimates? Unbiased! **BLUE**
    - **Incorrect sign errors** (type S): some counties really have higher kidney cancer rate than others?
    - **Incorrect magnitude** (type M): some counties really have higher/lower cancer rate?
  - Would correction for multiplicity help at all?
    - Useless in controlling for type S and M errors
- **How can we do better?**
  - Information share: across spatial elements
  - **Research hypothesis: P (effect > 0 | data)**

What do we know about spatial elements?

- **Massively univariate modeling**
  - Pretend full ignorance: fully trust the data
  - Uniform distribution: each element equally likely to have any value in (- ∞, +∞)
  - Similar for variances: variances can be negative in ANOVA
- **One crucial prior for spatial elements**
  - Reasonable to assume Gaussian distribution?
  - Gaussian assumption adopted everywhere!
    - Subjects, residuals across TRs
  - How can Gaussian assumption help?
    - Loosely constraining elements
    - No full trust for individual estimates
    - Information sharing: shrinkage or partial pooling
    - Controlling type S and M errors



Short summary: what we intend to achieve

- **Abandon strawman and *p*-value**
  - Directly focus on research interest: $P(\text{effect} > 0 \mid \text{data})$ vs. $P(\text{data} \mid \text{effect} = 0)$
- **Build one model**
  - Incorporate all elements into a multilevel or hierarchical structure
  - Loosely constrain elements: leverage **prior knowledge**
  - Achieve higher modeling efficiency: **no more multiplicity**!
  - Validate the model by comparing with potential competitors
  - Be conservative on effect estimates by controlling type S and M errors: **biased?**
  - Always be mindful of uncertainties: strength of evidence (no proof)
  - Less vulnerable to data manipulations: whole brain, gray matter, regions, …
- **Avoid dichotomous decisions**
  - Report full results if possible
  - Highlight instead of hide based on gradient of evidence
  - Focus on estimation, not inferences

Bayesian strategy in handling multiplicity

- **Conventional approach: neighborhood leverage**
  - Local relatedness: all regions act freely from each other
- **BML approach: global calibration**
  - Tug of war: **local effect vs global effect**
  - Weighted average
  - Partial pooling, shrinkage

$$\hat{\theta}_j = \frac{\frac{n}{\lambda^2+\sigma^2}\bar{y}_{\cdot j} + \frac{1}{\tau^2} b_0}{\frac{n}{\lambda^2+\sigma^2} + \frac{1}{\tau^2}}, \; V = \frac{1}{\frac{n}{\lambda^2+\sigma^2} + \frac{1}{\tau^2}}$$

Application: region-based analysis

- **Dataset**
  - Subjects: n = 124 children; resting-state data (Xiao et al., 2019)
  - Individual subjects: seed-based correlation for each subject
    - 3D correlation between seed and whole brain ("functional connectivity")
  - Explanatory variable (behavior data): Theory of Mind Index x_i

Uniform distribution: total freedom - each parameter on its own

- **Voxel-wise group analysis: GLMs**
  - Focus: association between x and seed-based correlation (z-score)
  - Pretense: voxels unrelated - equal likelihood within (-∞, ∞)
  - Information waste!
  - GLMs: mass univariate - multiplicity

  m = 100,000 voxels → 100,000 models

$$
\begin{aligned}
\text{1st voxel: } \boldsymbol{y}_1 &= a_1 + b_1\boldsymbol{x} + \boldsymbol{\epsilon}_1 \\
\text{2nd voxel: } \boldsymbol{y}_2 &= a_2 + b_2\boldsymbol{x} + \boldsymbol{\epsilon}_2 \\
&\cdots \\
m\text{th voxel: } \boldsymbol{y}_m &= a_m + b_m\boldsymbol{x} + \boldsymbol{\epsilon}_m
\end{aligned}
$$

Xiao et al., 2019. Neuroimage 184:707-716

GLMs: dealing with multiplicity!

- **Voxel-based analysis: GLMs**
  - Penalty time for pretense: multiple testing (m = 100,000), magic 0.05
  - Show time for various correction methods
    - Voxel-wise p, FWE, FDR, spatial smoothness, clusters, …
    - Simulations, random field theory, permutations, …
    - How would dataset turn out under GLM? 4 lucky clusters managed to survive

voxel p	cluster threshold	surviving ROIs	ROIs
0.001	28	2	R PCC, PCC/PrC
0.005	66	4	R PCC, PCC/PrC., L IPL, L TPJ
0.01	106	4	R PCC, PCC/PrC., L IPL, L TPJ
0.05	467	4	R PCC, PCC/PrC., L IPL, L TPJ

Switching from voxels to ROIs: still GLMs

- **Region-wise analysis : GLMs**
  - Focus: association between and seed-based correlation (z-score)
  - Pretense: ROIs unrelated
  - GLMs: mass univariate
    m = 21 ROIs ⟶
    21 models
  - Penalty for pretense:
    multiple testing – what to do?
    - Bonferroni? Unbearable
    - What else?

Uniform distribution: total freedom - each parameter on its own.

$$\begin{aligned}
\text{1st ROI: } \boldsymbol{y}_1 &= a_1 + b_1\boldsymbol{x} + \boldsymbol{\epsilon}_1 \\
\text{2nd ROI: } \boldsymbol{y}_2 &= a_2 + b_2\boldsymbol{x} + \boldsymbol{\epsilon}_2 \\
&\cdots \\
m\text{th ROI: } \boldsymbol{y}_m &= a_m + b_m\boldsymbol{x} + \boldsymbol{\epsilon}_m
\end{aligned}$$

Switching from GLMs to LME

- **Region-wise analysis : Linear Mixed-Effects (LME) model**
  - One model integrates all regions
  - ROIs loosely constrained instead of being unrelated
    - Gaussian distribution: Is it far-fetched or subjective?
    - Similar to cross-subject variability
  - Goal: effect of interest- $a + \alpha_j$, $b + \beta_j$
  - Differentiation: fixed vs. random
    - Fixed: **epistemic** uncertainty
    - Random: **aleatoric** uncertainty
    - Julius Caesar: Alea iacta est. January 10, 49 BC
  - What can we get out of LME?
    - Conventional framework
    - Estimates for fixed effects
    - Variances for random effects
  - ***Dead end!***

New components

Overall effect: shared by all ROIs and subjects

idiosyncratic effect of *i*th subject

Unique effect of *j*th ROI

$$z_{ij} = \boxed{a + bx_i} + \boxed{\pi_i} + \boxed{\alpha_j + \beta_j x_i} + \epsilon_{ij}$$

$$\pi_i \overset{iid}{\sim} \mathcal{N}(0, \tau^2), (\alpha_j, \beta_j)^T \overset{iid}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\lambda})$$

$$\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2), i = 1, 2, ..., n, \ j = 1, 2, ..., m$$

Switching from GLMs to BML

- **Region-wise analysis : Bayesian multilevel (BML) model**
  - One model integrates all regions: basically same as LME
  - ROIs loosely constrained instead of being unrelated
    - Gaussian distribution: Is it far-fetched or subjective?
    - Similar to cross-subject variability
  - Goal: effect of interest $b + \beta_j$
  - No more differentiation: fixed vs. random
    - All parameters: **aleatoric**
  - Same model as LME plus priors
    - Markov chain Monte Carlo (MCMC)
    - Inferences via posterior distribution
  - ***Ka-ching!***

New components

Overall effect: shared by all ROIs and subjects

Idiosyncratic effect by *i*th subject

Unique effect by *j*th ROI

$$z_{ij} = \boxed{a + bx_i} + \boxed{\pi_i} + \boxed{\alpha_j + \beta_j x_i} + \epsilon_{ij}$$

$$\pi_i \overset{iid}{\sim} \mathcal{N}(0, \tau^2), (\alpha_j, \beta_j)^T \overset{iid}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\lambda})$$

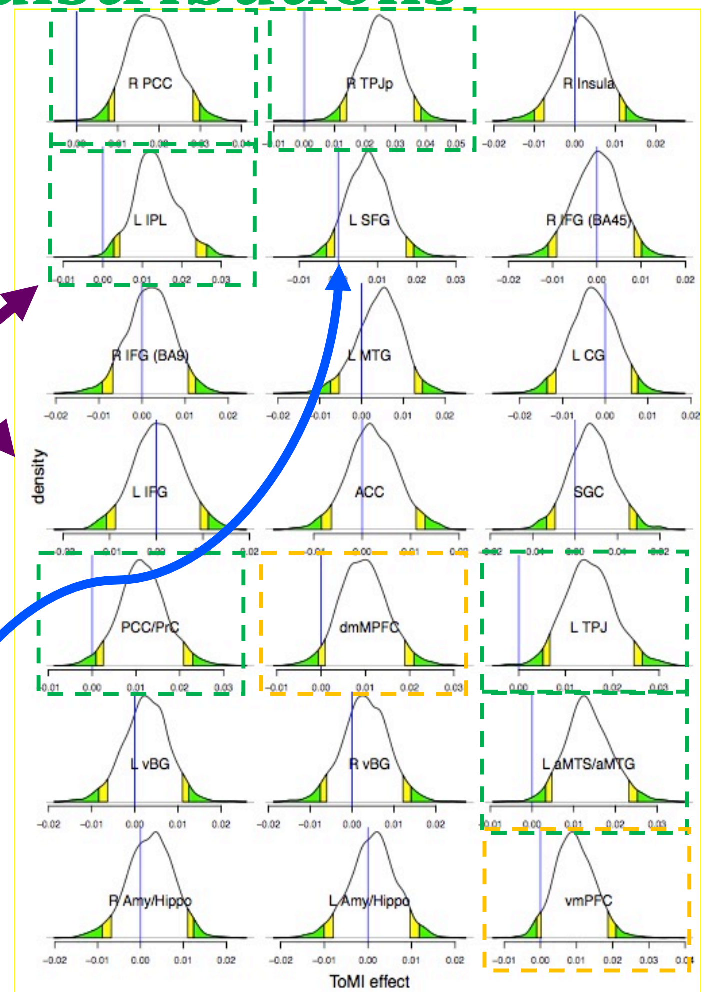
$$\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2), i = 1, 2, ..., n, \ j = 1, 2, ..., m$$

Chen, et al, 2019. Handling Multiplicity in Neuroimaging through Bayesian Lenses with Multilevel Modeling. Neuroinformatics.

Inferences from BML: full distributions

- **Region-based BML: 21 ROIs**
- **Full report with richer information: posterior distributions for each ROI**
    - No dichotomization
    - No results hiding
    - No discrimination against small regions
    - No ambiguities about spatial specificity
    - No inconvenient interpretation of confidence interval
    - Evidence for each ROI: ***P*** **(effect > 0 | data)**

Highlight, not hide



- **9 ROIs with strong evidence of effect compared to**
    - Region-wise GLM with Bonferroni correction
    - Voxel-wise GLM at cluster level: 2 clusters

How about Left SFG?

Inferences from BML: uncertainty

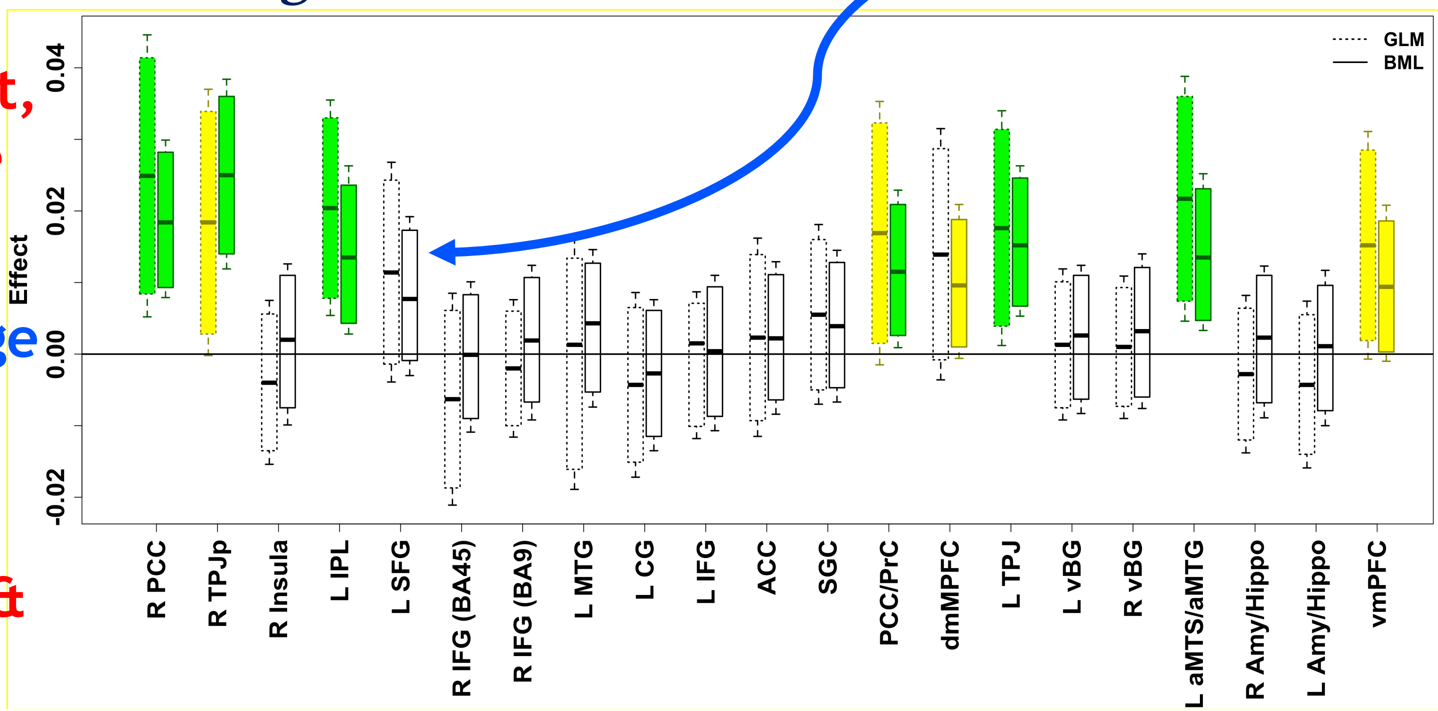
- ROI-based BML: 21 ROIs
- Full report with bar graph uncertainty intervals
  - Nothing hidden under sea level
- 8 ROIs with strong evidence for effect of interest

How about Left SFG?

Highlight, not hide

Shrinkage / partial pooling

Type M & S errors



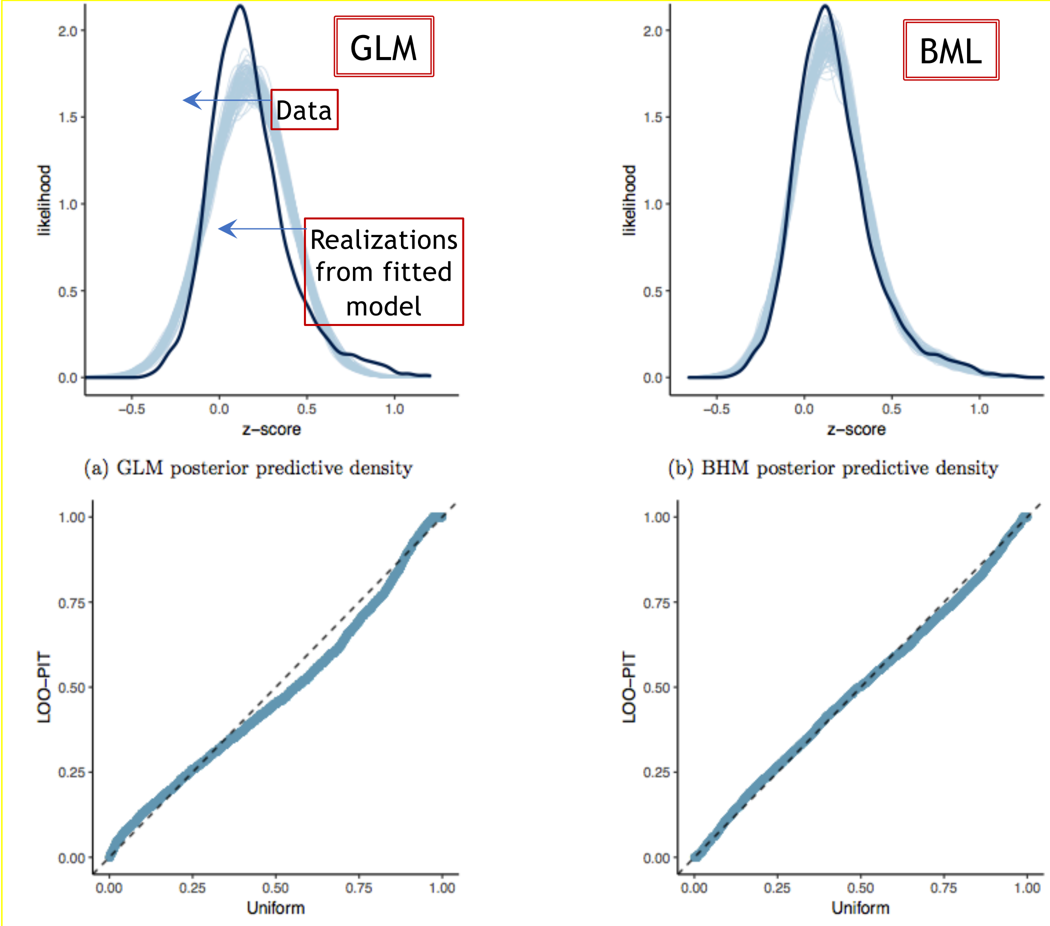
BML: model validations

- **Cross-validation**
  - Leave-one-out information criterion (LOOIC)

Cross-validation

	LOOIC	SE
GLM	-300.39	98.25
BML	-2247.06	86.42
GLM - BML	1946.67	96.35

  - Posterior predictive checking
- **Effects of BML**
  - Regularizing ROIs: don't fully trust individual ROI data
  - Sacrificing fit at each ROI; achieving better overall fit



(a) GLM posterior predictive density

(b) BHM posterior predictive density

(c) GLM cross-validation: Q-Q plot (uniform)

(d) BHM cross-validation: Q-Q plot (uniform)

BML: Whole-brain vs. region-base analysis

- **Region-based analysis**
  + high region specificity: region definitions considered as priors
  + low computational cost
  + avoiding potential alignment issues by defining regions in native space: FreeSurfer + SUMA
  - not all regions have been defined
  - information loss due to averaging within each region
  - region definitions can be tricky
    - relying on results accuracy in literature (e.g., publication bias)
    - different atlases/parcellations

- **Whole-brain analysis**
  + independent of region definitions
  + less likely to miss small regions that are not in available atlases/parcellations
  - vulnerable to poor alignment across subjects
  - region specificity problem
    - Voxel-wise results do not respect region definitions
  - Computationally challenging
    - hopeful: within-chain parallelization and GPU usage

Application #2: matrix-based analysis

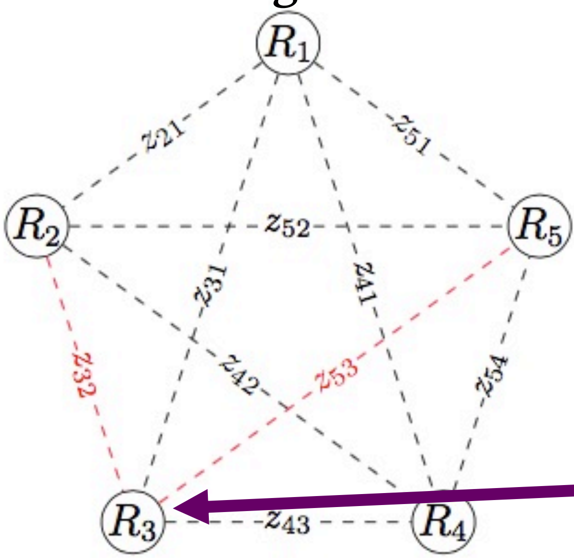
- **Dataset: correlation matrix**
  - Subjects: n = 41 subjects; response-conflict task (Choi et al., 2012)
  - Individual subjects: correlation matrix among m = 16 ROIs
  - How to go about group analysis?
    - **GLM for each element in correlation matrix: NBS, CONN, FSLNets in FSL, GIFT**
    - **Binarization approach: graph theory**
  - More broadly: matrix-based analysis (MBA) ("network modeling")
    - **Inter-region correlation (IRC): FMRI**
    - **White matter properties (FA, MD, …): DTI**
    - **Other matrices (e.g., coherence, entropy, mutual information)**
- **Focus on GLM**
  - Student t-test or GLM on each element
    - M = 120 massively univariate models
  - Pretense again: all elements are unrelated
  - Equal likelihood within $(-\infty, \infty)$
  - Information waste
  - Penalty time again: permutations? FDR?

$$\boldsymbol{Z}_k^{(m)} = \begin{array}{cc} & \begin{array}{ccccc} R_1 & R_2 & R_3 & \cdots & R_m \end{array} \\ \begin{array}{c} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_m \end{array} & \begin{pmatrix} - & z_{12k} & z_{13k} & \cdots & z_{1mk} \\ z_{21k} & - & z_{23k} & \cdots & z_{2mk} \\ z_{31k} & z_{32k} & - & \cdots & z_{3mk} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{m1k} & z_{m2k} & z_{m3k} & \cdots & - \end{pmatrix} \end{array}$$

Choi et al., 2012. Neuroimage 59(2):1912-1923

Dealing with inter-region correlations (IRCs)

- **Complexities of IRCs**
  - Some region pairs are unrelated, but others are correlated
  - Correlation structure is intricate
  - $0 \leq \rho \leq 0.5$
  - Can we do a better job than GLMs or dichotomization?
    - **Challenge: How to characterize the complex structure?**



$$\boldsymbol{P}^{(5)} = \begin{array}{c} \\ z_{21} \\ z_{31} \\ z_{41} \\ z_{51} \\ z_{32} \\ z_{42} \\ z_{52} \\ z_{43} \\ z_{53} \\ z_{54} \end{array} \begin{array}{c} \begin{array}{cccccccccc} z_{21} & z_{31} & z_{41} & z_{51} & z_{32} & z_{42} & z_{52} & z_{43} & z_{53} & z_{54} \end{array} \\ \begin{pmatrix} 1 & \rho & \rho & \rho & \rho & \rho & \rho & 0 & 0 & 0 \\ \rho & 1 & \rho & \rho & \rho & 0 & 0 & \rho & \rho & 0 \\ \rho & \rho & 1 & \rho & 0 & \rho & 0 & \rho & 0 & \rho \\ \rho & \rho & \rho & 1 & 0 & 0 & \rho & 0 & \rho & \rho \\ \rho & \rho & 0 & 0 & 1 & \rho & \rho & \rho & \rho & 0 \\ \rho & 0 & \rho & 0 & \rho & 1 & \rho & \rho & 0 & \rho \\ \rho & 0 & 0 & \rho & \rho & \rho & 1 & 0 & \rho & \rho \\ 0 & \rho & \rho & 0 & \rho & \rho & 0 & 1 & \rho & \rho \\ 0 & \rho & 0 & \rho & \rho & 0 & \rho & \rho & 1 & \rho \\ 0 & 0 & \rho & \rho & 0 & \rho & \rho & \rho & \rho & 1 \end{pmatrix} \end{array}$$

IRC: switching from GLM to LME

- **IRC analysis through linear mixed-effects (LME) modeling**
  - One model integrates all ROIs: LME
  - ROIs loosely constrained instead of being unrelated
    - Gaussian distribution: Is it far-fetched?
    - Similar to cross-subject variability
  - Differentiation: fixed vs. random
    - Fixed: **epistemic** uncertainty
    - Random: **aleatoric** uncertainty
  - Effects of interest
    - region pair: $b_0+\xi_i+\xi_j+\eta_{jj}$
    - region: $0.5*b_0+\xi_i$
  - **LME wouldn't work!**
    Dead end!

Unique effect at *i*th & *j*th ROI

Unique effect of RP

Unique effect at *i*th & *j*th ROI for *k*th subject

overall effect: shared by all ROIs and subjects

unique effect by *k*th subject

$$z_{ijk} = b_0 + \boxed{\xi_i + \xi_j + \eta_{ij} + \zeta_{ik} + \zeta_{jk} + \pi_k} + \epsilon_{ijk}$$

$$\xi_i, \xi_j \stackrel{iid}{\sim} \mathcal{N}(0,\lambda^2),\ \eta_{ij} \sim \mathcal{N}(0,\mu^2),\ \zeta_{ik}, \zeta_{jk} \sim \mathcal{N}(0,\nu^2),\ \pi_k \stackrel{iid}{\sim} \mathcal{N}(0,\tau^2),\ \epsilon_{ijk} \sim \mathcal{N}(0,\sigma^2)$$

$$i, j = 1, 2, ..., m\ (i > j),\ k = 1, 2, ..., n$$

IRC: one more jump from LME to BML

- **IRC analysis through Bayesian multilevel (BML) modeling**
  - One model integrates all ROIs: BML (essentially same as LME)
  - ROIs loosely constrained instead of being unrelated
    - Gaussian distribution: Is it far-fetched?
    - Similar to cross-subject variability
  - No differentiation: fixed vs. random
    - All parameters: **aleatoric** uncertainty
  - Effects of interest
    - region pair: $b_o+\xi_i+\xi_j+\eta_{jj}$
    - region: $0.5{*}b_o + \xi_i$
  - LME plus priors
    - MCMC
    - Posterior distribution
  - ***Ka-ching!***

Unique effect at *i*th & *j*th ROI

unique effect of RP

Unique effect at *i*th & *j*th ROI for *k*th subject

overall effect: shared by all ROIs and subjects

unique effect by *k*th subject

$$z_{ijk} = b_0 + \boxed{\xi_i + \xi_j + \eta_{ij} + \zeta_{ik} + \zeta_{jk} + \pi_k} + \epsilon_{ijk}$$

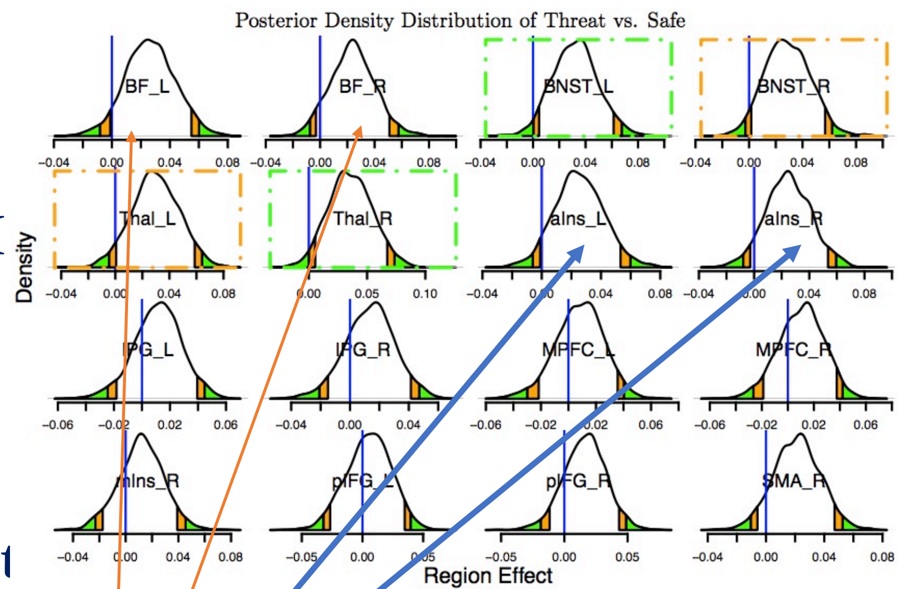
$$\xi_i, \xi_j \overset{iid}{\sim} \mathcal{N}(0,\lambda^2),\ \eta_{ij} \sim \mathcal{N}(0,\mu^2),\ \zeta_{ik},\zeta_{jk} \sim \mathcal{N}(0,\nu^2),\ \pi_k \overset{iid}{\sim} \mathcal{N}(0,\tau^2),\ \epsilon_{ijk} \sim \mathcal{N}(0,\sigma^2)$$

$$i,j = 1,2,\ldots,m\ (i>j),\ \ k = 1,2,\ldots,n$$

Chen, et al, 2019. An integrative Bayesian approach to matrix-based analysis in neuroimaging. Human Brain Mapping.

IRC – ROI effect from BML: full distributions

- **ROI-based BML: 16 ROIs**
- **Full report with richer information: posterior distributions for each ROI**
  - No dichotomization
  - Nothing hidden under sea level
  - $\rho = 0.483$
- **4 ROIs with strong evidence of effect compared to**
  - Region effect inferences: unavailable from GLM and graph theory
  - Hubness?

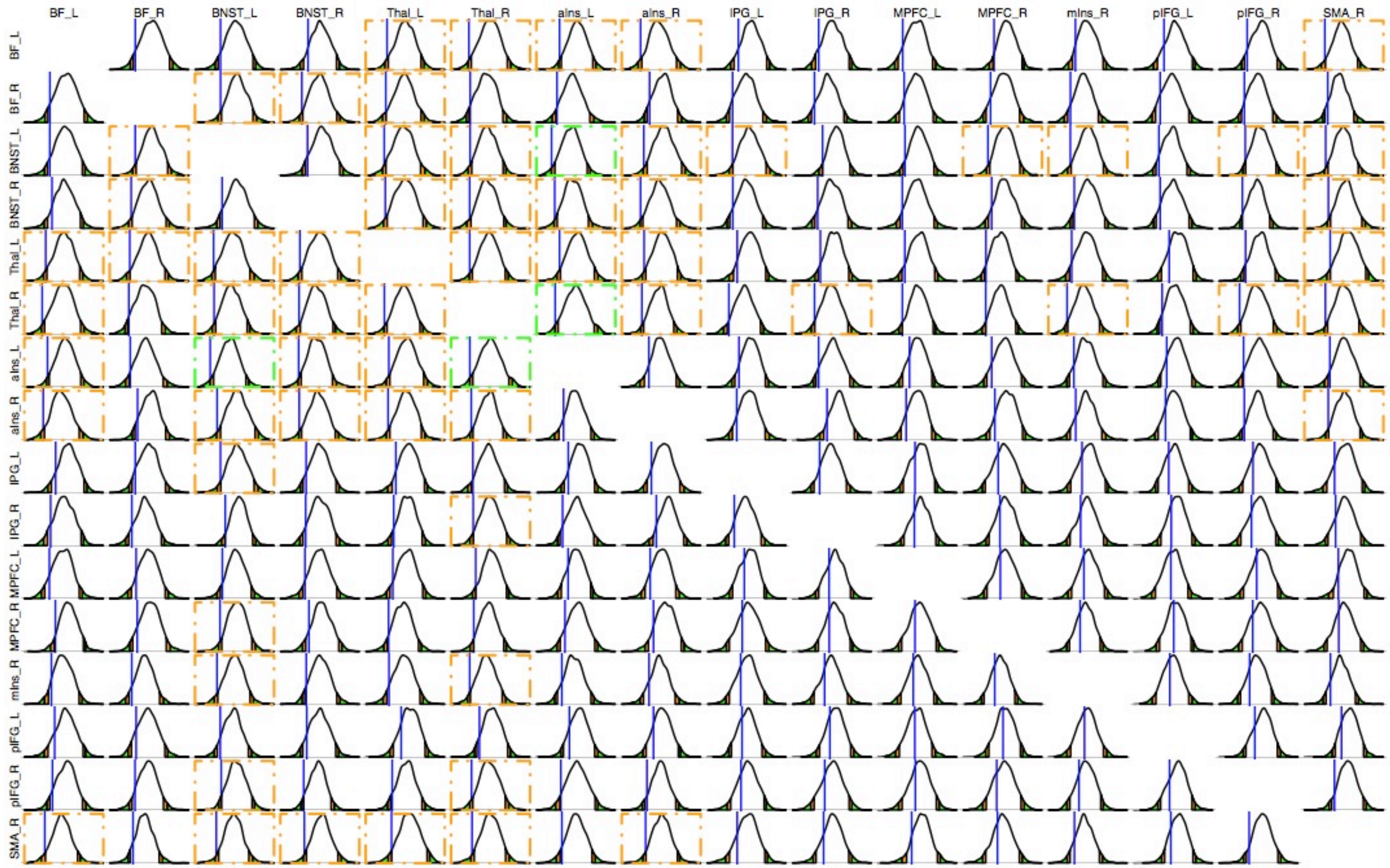


How about basal forebrain and Anterior Insula: L & R?

Highlight, not hide

IRC – RP effect from BML: full distributions

120 RPs

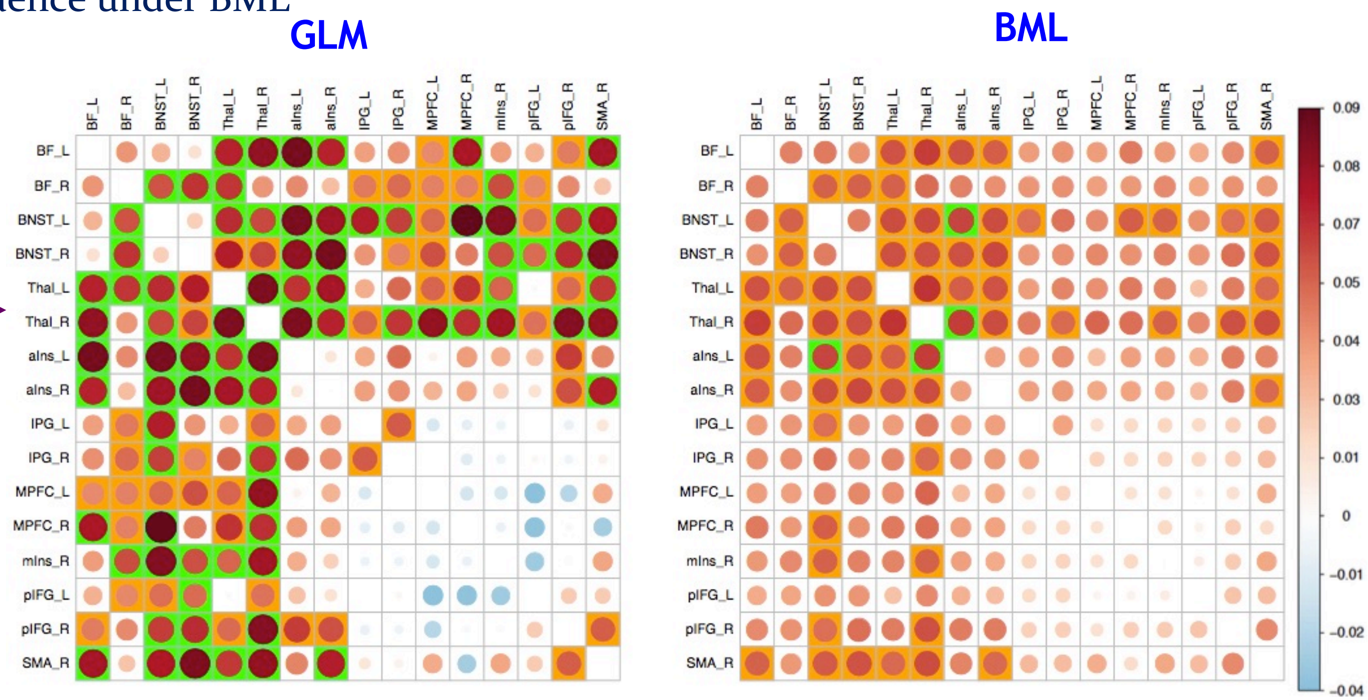


Highlight, not hide

IRC– RP effect from BML

- **ROI-based BML: 16 ROIs**
- **Full report for all region pairs (RPs)**
- **Comparisons with GLMs: nothing hidden under sea level**
  - 63 RPs identified by GLMs with *p* of 0.05: none survived after correction with NBS via permutations
  - 33 RPs with strong evidence under BML

Highlight, not hide



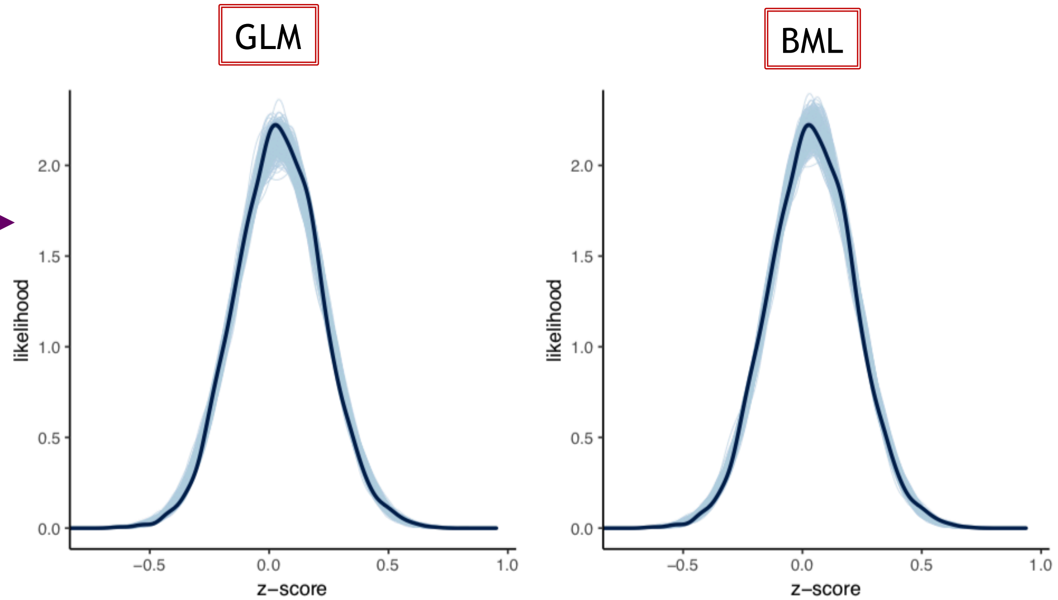
BML: model validations

- **ROI-based BML with IRD of 16 ROIs: cross-validation**
  - Leave-one-out information criterion (LOOIC)

Cross-validation

Model	LOOIC	SE
GLM	-2808.31	101.65
BML0	-4543.77	102.97

  - Posterior predictive checking



- **Effects of BML**
  - Regularizing ROIs: don't fully trust individual ROI data
  - Sacrificing fit at each ROI; achieving better overall fit

Bayesian all the way

- **Should one correct for a duplicated study?**
- **How about all studies with statistical analyses**
- **Everyone is Bayesian**
  - Probabilistic nature
    - data; preprocessing, subjects, groups, sites, scanners, modeling approaches
  - Reproducibility
    - Most studies: similar; minority: outliers
    - Applying a Gaussian prior
- **Embracing, not fighting, multiplicity!**

Contrast

- **Mass Univariate Approach**
  - Accurate with the current data, but poor for predictions
  - Trust effect estimates (unbiased), but don’t report them
  - Doubt about statistical evidence, but selectively report it through filtering with colorbar and in table

- **BML**
  - Compromise with the current data, but gain accuracy for predictions
  - Pool effect estimate toward the center (biased), and directly show them through posterior distributions
  - Statistical evidence shown without filtering

Summary

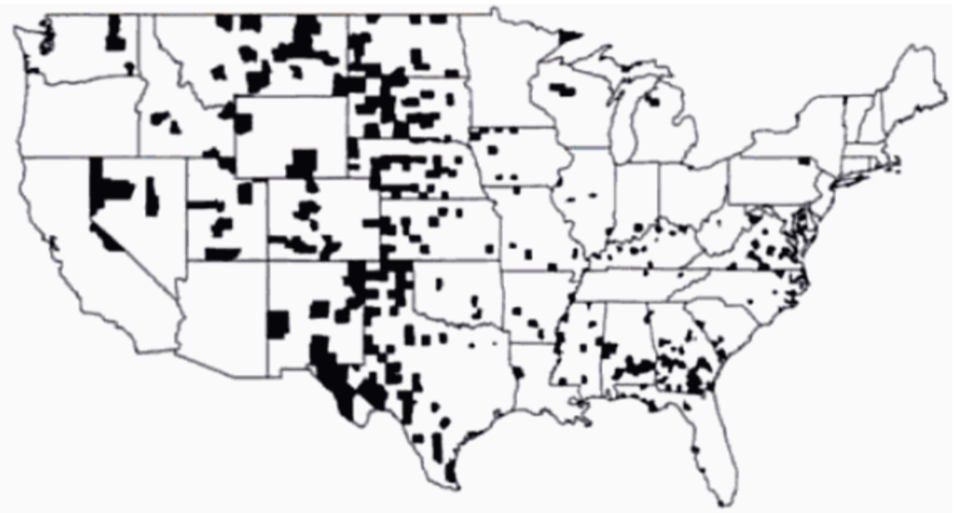
- **Issues with current correction for multiplicity**
- **Two toy examples**
  - NBA players
  - Kidney cancer
- **Application: region-based analysis (RBA)**
  - Program in AFNI: ***RBA***
- **Other applications**
  - Matrix-based analysis (program in AFNI: ***MBA***)
  - Region-based inter-subject correlation (ISC) analysis
  - Gray matter connectivity analysis
  - Others cases involving multiplicity

Keep Kidney Cancer in Mind!

- **Kidney cancer distribution among counties**

Highest rate

lowest rate



Calibration, regularization, information sharing, partial pooling, shrinkage

Acknowledgements

- **Paul-Christian Bürkner** (Aalto University, Finland)
- **Andrew Gelman** (Columbia University), **Stan Development Team, R Foundation**

- **Yaqiong Xiao, Elizabeth Redcay, Tracy Riggins, Fengji Geng**
- **Luiz Pessoa, Joshua Kinnison** (Depart of Psychology, University of Maryland)

- **Zhihao Li** (School of Psychology and Sociology, Shenzhen University, China)
  Lijun Yin (Department of Psychology, Sun Yat-sen University, China)

- **Emily Finn, Daniel Handwerker** (SFIM/NIMH, National Institutes of Health)

- **Paul A. Taylor, Daniel R. Glen, Justin K. Rajendra, Richard C. Reynolds, John Lee, Robert W. Cox** (SSCC/NIMH, National Institutes of Health)