



Univariate vs. Multivariate Lesion Symptom Mapping Approaches

J. Baldo¹, M. Ivanova^{1,2,3}, B. Curran^{1,4}, N. Dronkers^{1,2} & T. Herron¹



¹VA Northern California Health Care System, ²Univ of California, Berkeley, ³National Research University, HSE, Moscow, Russia, ⁴Univ of California, San Francisco

Introduction

Mass-univariate lesion symptom mapping methods (ULSM), such as the original voxel-based lesion symptom mapping (VLSM), provide statistical comparisons of behavioral performance in brain-injured patients with and without a lesion on a voxel by voxel basis^{1,2}. New multivariate lesion-symptom mapping (MLSM) methods have been developed that consider the entirety of all lesion patterns (all measurement units) simultaneously in one model^{3,4}. Advantages and disadvantages of ULSM and MLSM techniques have been discussed in the literature, but very little work has been done to empirically test specific claims.⁴ In the current study, we directly compared ULSM and MLSM methods by analyzing their performance on both artificial and real datasets of brain-behavioral relationships (BBRs).

Procedures: Synthetic Data & Real Data

Synthetic Data:

1. Single parcel, proportional BBR conditions (% of target lesioned ~ % of behavioral deficit) in the left middle cerebral artery (MCA) territory over a fully crossed design:

- 16 or 30 GM parcels of Left MCA as BBR targets
- lesion masks from our site (n=209) and another site⁴ (n= 131)
- 13 lesion symptom mapping methods (8 MLSM)
- 4mm lesion mask smoothing vs. none
- 7 different patient sample sizes: n=32,48,64,80,96,112, &128
- multiple spatial accuracy measures (6 distance & 2 overlap)
- 3 behavioral noise levels

2. Procedure above was repeated with two-parcel networks, testing redundant, dependent, and extended networks.

Real Data:

Western Aphasia Battery language data from LH stroke patients:

- Overall Aphasia Score Improvement Over Time.
- Auditory Comprehension Subscore Improvement.
- Single Word Comprehension Subscore Improvement.

Procedures: LSM Methods Tested

Multivariate LSM*

ICA-L1 ICA - Independent component analysis
 ICA-L2 ICA - Logistic principal component analysis
 LPCA-L1 LPCA - Logistic principal component analysis
 LPCA-L2 LPCA - Logistic principal component analysis
 SVD-L1 SVD - Singular value decomposition
 SVD-L2 SVD - Singular value decomposition
 PLS Partial least squares (dense)
 SVR Support Vector Regression⁵

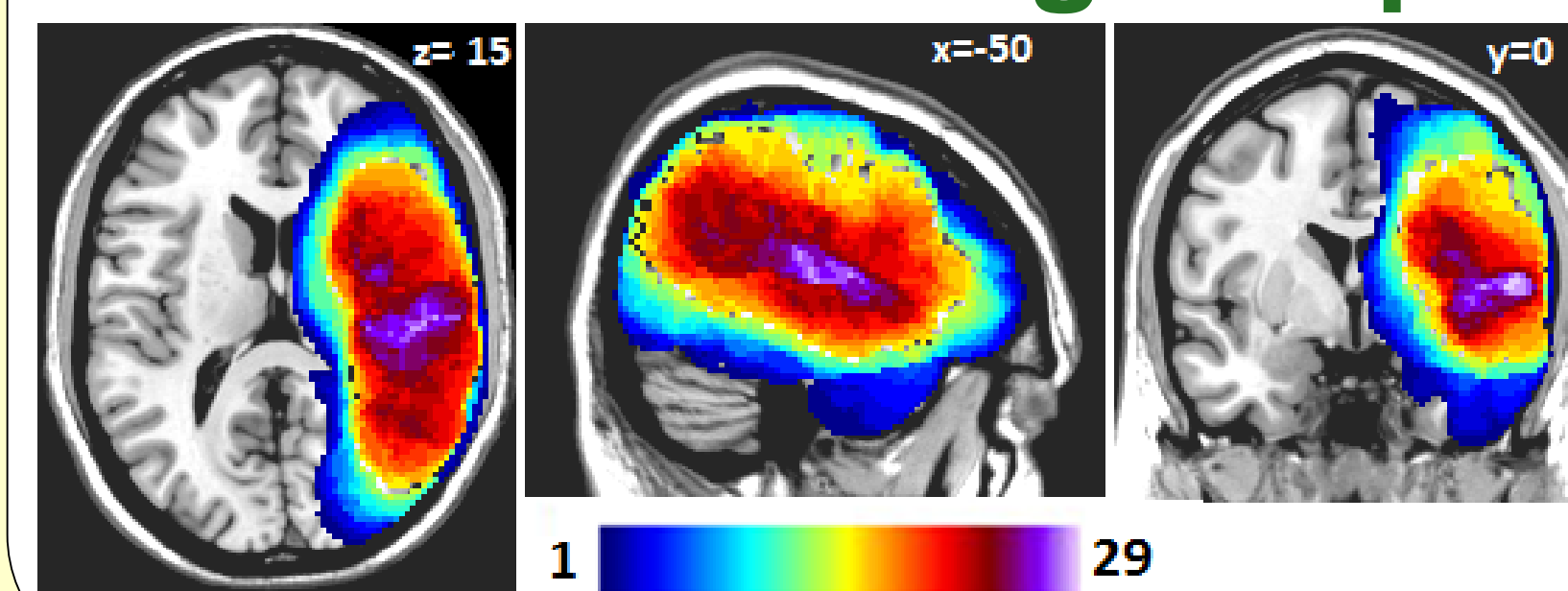
*[L1 - elastic net regression ; 95% L1 penalty]
[L2 - elastic net regression ; 95% L2 penalty]

Univariate LSM**6

T-max Maximum t value
 T-nu=125 125th highest t value [Mirman]
 T-0.0001 cluster size when p<0.0001
 T-0.001 cluster size when p<0.001
 T-0.01 cluster size when p<0.01

**All U-VLSM methods used linear regression at every voxel plus permutation testing to set familywise thresholds based on five different criteria listed above.

Lesion Coverage Map



Overlay of stroke patients' lesions from our site, showing voxels included in the real LSM analyses. Color bar shows the degree of lesion overlap.

- 46 single chronic left stroke (9 female)
- mean post-stroke: 42 months (range 1-328)
- mean age: 61 (range 31-86)
- All subjects initial Comprehension subscore was <9 at initial WAB measurement.

Results: Single Anatomical BBR Target

A	32	48	64	80	96	112	128
ICA-L1	0.74	0.89	0.97	0.99	0.99	1	1
ICA-L2	0.77	0.91	0.98	1	1	1	1
LPCA-L1	0.80	0.93	0.99	0.99	1	1	1
LPCA-L2	0.81	0.95	0.99	1	1	1	1
PLS	0.69	0.88	0.97	0.99	1	1	1
SVD-L1	0.76	0.88	0.96	0.99	1	1	1
SVD-L2	0.75	0.88	0.94	0.98	1	1	1
SVR	0.85	0.96	1	1	1	1	1
T-max	0.91	0.99	1	1	1	1	1
T-0.0001	0.91	0.99	1	1	1	1	1
T-0.001	0.93	0.99	1	1	1	1	1
T-0.01	0.89	0.98	1	1	1	1	1

B	32	48	64	80	96	112	128
0.00	0.95	0.99	1	1	1	1	1
0.38	0.88	0.98	1	1	1	1	1
0.77	0.67	0.84	0.96	0.98	0.99	1	1

LSM Power: Fraction of time that LSM produces a cluster ostensibly identifying the target.
A: LSM method (rows) vs. # of Patients in LSM (columns).
B: Behavioral Noise Level (rows; fraction of behavioral std. dev. White noise added) vs. # of Patients (columns). ICA, SVD, LPCA are the lesion mask data reduction methods.

LSM Method	mm	Target Center	mm	# of Patients	mm
ICA-L1	6.4	COM	5.9	32	5.4
ICA-L2	7.1	AnyHit	3.1	48	4.7
LPCA-L1	6.0			64	4.5
LPCA-L2	5.9			80	4.4
PLS	6.4	Mask Smoothing	mm	96	4.2
SVD-L1	4.0	0mm	4.6	112	4.2
SVD-L2	3.9	4mm	4.3	128	4.2
SVR	2.8				
T-max	2.8				
T-0.0001	3.0	Cluster Location	mm	Noise Level	Mm
T-0.001	3.3	COM	5.4	0.00	4.5
T-0.01	3.6	Max	3.5	0.38	4.5
T-nu=125	3.3	wCOM	4.6	0.77	4.4

LSM Accuracy: Distance from LSM Cluster center to Anatomical Target center averaging over multiple center definitions. **Target Center:** COM: Center of Mass ; **AnyHit:** closest target location. **Cluster Location:** COM: cluster Center Of Mass. **Max:** maximum LSM statistic voxel location. **wCOM:** Weighted cluster center of mass. **Mask Smoothing:** Gaussian smoothing (FWHM). **Noise Level:** see LSM Power Table.

Results: Two Parcel BBR Target (Network)

	64	80	96	112	128
Fragile	0.99	0.99	1.00	1.00	1.00
Extended	0.98	0.99	1.00	1.00	1.00
Redundant	0.81	0.88	0.91	0.94	0.96

LSM Power: LSM success fraction: # of Subjects (columns) vs. two anatomical target network type (rows).

LSM Methods	Dice	# of patients	Dice
ICA-L1	0.09	64	0.11
ICA-L2	0.07	80	0.1
LPCA-L1	0.07	96	0.1
LPCA-L2	0.07	112	0.1
PLS	0.06	128	0.09
SVD-L1	0.08		
SVD-L2	0.09		
SVR	0.15		
T-max	0.15		
T-0.0001	0.14		
T-0.001	0.12		
T-0.01	0.1		
T-nu=125	0.13		

Network Type	Dice
Fragile	0.09
Extended	0.1
Redundant	0.11

LSM Accuracy: Dice coefficients for above threshold LSM clusters vs. two target network.

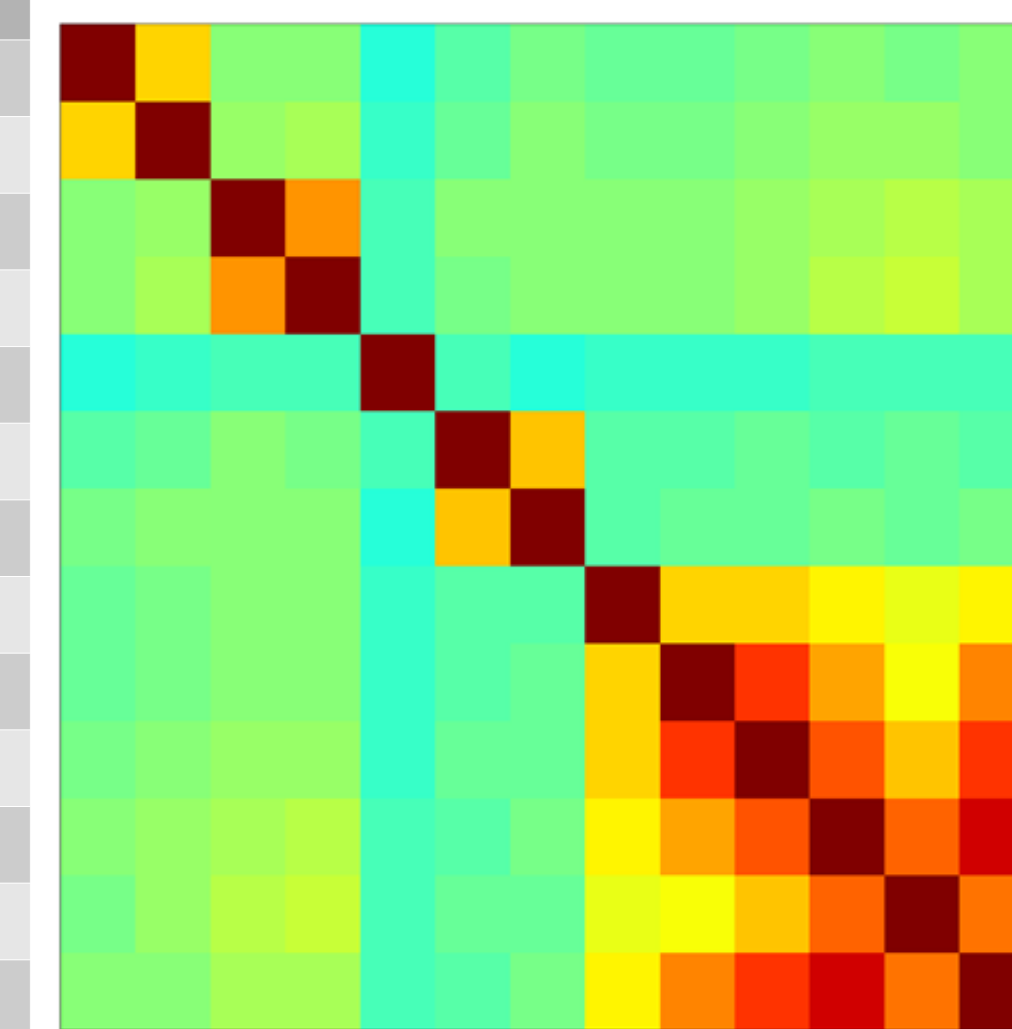
	64	80	96	112	128
Fragile	-0.20	-0.07	+0.01	+0.08	+0.13
Extended	-0.13	-0.00	+0.10	+0.16	+0.22
Redundant	-0.35	-0.25	-0.17	-0.09	-0.04

LSM Accuracy: Distribution comparisons of LSM values inside targets vs. outside targets using a one-sided Kuiper test [+1=best, -1=worst] for Network type and sample size (above) or LSM methods and sample size (below)

	64	80	96	112	128
ICA-L1	-0.44	-0.35	-0.26	-0.18	-0.13
ICA-L2	-0.24	-0.15	-0.04	0.04	0.1
LPCA-L1	-0.06	0.08	0.19	0.25	0.32
LPCA-L2	-0.01	0.12	0.23	0.3	0.36
PLS	-0.09	0.03	0.08	0.12	0.16
SVD-L1	0	0.15	0.25	0.29	0.32
SVD-L2	-0.01	0.16	0.22	0.3	0.33
SVR	-0.64	-0.53	-0.43	-0.36	-0.28
T-max	-0.52	-0.41	-0.31	-0.24	-0.17
T-0.0001	-0.37	-0.26	-0.17	-0.09	-0.03
T-0.001	-0.21	-0.11	-0.02	0.04	0.09
T-0.01	-0.03	0.07	0.14	0.19	0.22
T-nu=125	-0.3	-0.19	-0.1	-0.04	0.02

Results: Zero Targets (False Positives)

Voxels	Clusters	
715	4.7	ICA-L1
719	5.3	ICA-L2
1620	5.6	LPCA-L1
1877	5.3	LPCA-L2
2435	5.8	PLS
873	6.9	SVD-L1
963	7.3	SVD-L2
17	1.5	SVR
17	1.5	t_max
73	1.2	t_0.0001
452	1.0	t_0.001
2324	1.0	t_0.01
312	4.5	t_nu=125



False Positive Properties: Table shows mean total number of voxels (size 8mm³) broken up into a (mean) number of contiguous clusters for each solution method type. Figure displays inter-method false positive correlations - i.e. how often two methods have co-occurring false positive LSM maps.

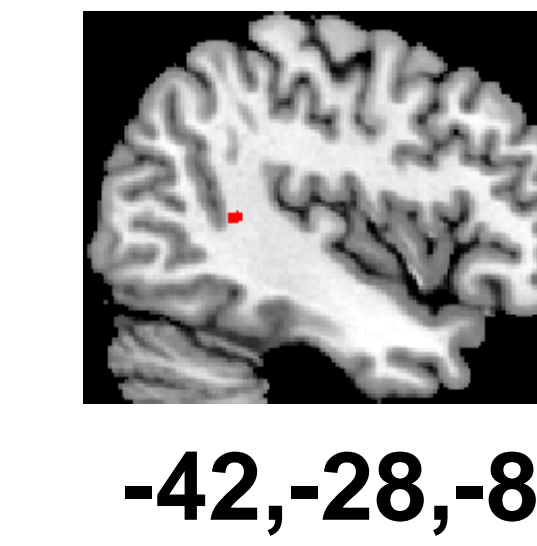
Results: LSM Output with Real Language Data

Aphasia Q.

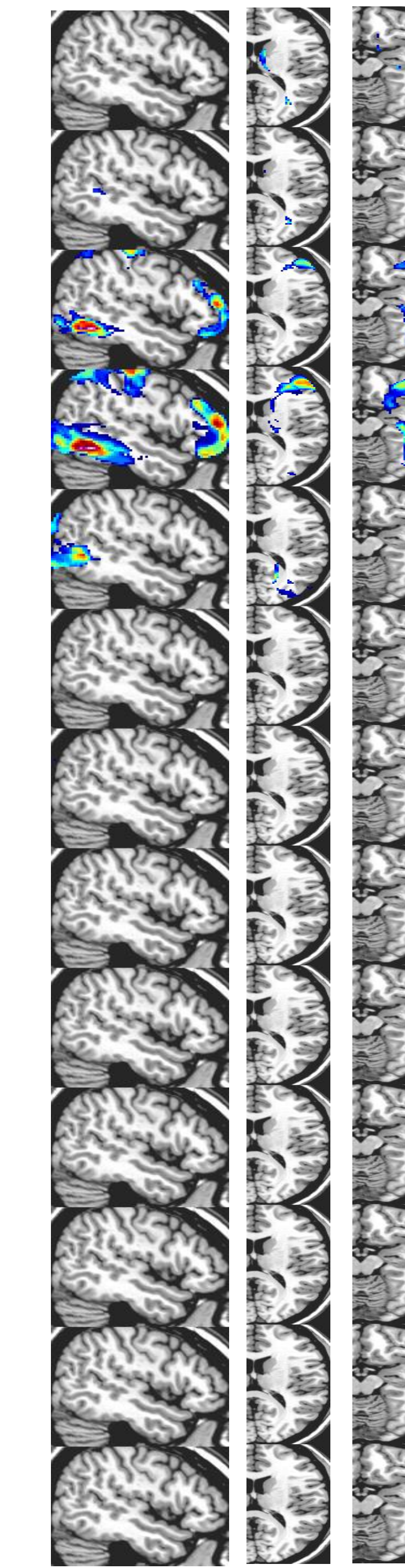
Aud Comp.

1 Word Comp

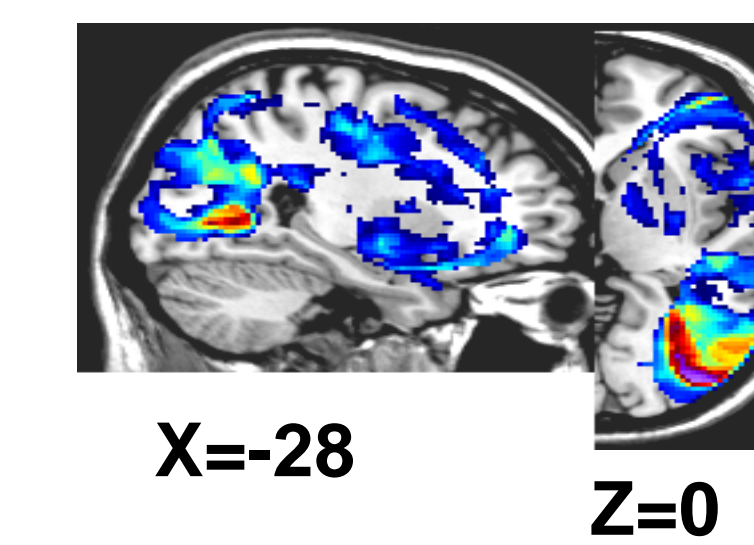
- ICA-L1
- ICA-L2
- LPCA-L1
- LPCA-L2
- PLS
- SVD-L1
- SVD-L2
- SVR
- T-max
- T-0.0001
- T-0.001
- T-0.01
- T-nu=125



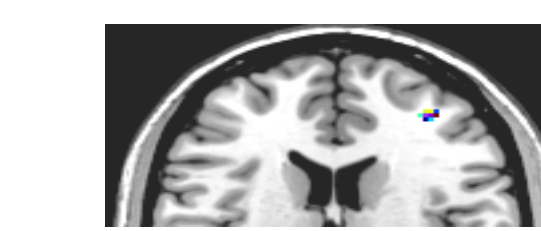
-42,-28,-8



X=-50 Z=+15 Z=-15



X=-28 Z=0



Z=+16

LSM results on real data for WAB Aphasia Quotient improvement, Auditory Comprehension improvement, and Single Word Comprehension improvement from multiple WABs per subject, covaried for lesion size, age, and gender. The minimum power per voxel was 0.1 at p<0.01 or 5 subjects w/ lesions.

Summary

Modern ULSM techniques⁶ provide a robust solution for detecting single targets, and required a smaller sample size than MLSM to achieve a similar level of power and spatial accuracy.

- With certain metrics, some (but not all) MLSM methods have advantages for detecting two-target networks, but cluster-size based ULSM methods can also provide insight into this case.
- Noise level has a modest impact on ULSM and MLSM results, mostly affecting LSM power.
- ULSM methods do better with noiseless data, but certain distance metrics reduce LSM cluster spatial sensitivity to behavioral noise.
- Smoothing at 4mm improves accuracy of localization across all metrics for both ULSM and MLSM methods, despite there being no anatomical imprecision in the synthetic models.
- Weighted center-of-mass (wCOM) and peak statistical value (Max) locations of obtained LSM clusters provide the most robust accuracy results across all methods.
- Dice overlap scores were unacceptably low for all methods (even for single targets), but distributional comparisons (inside vs. outside target(s)) proved useful for method evaluation.
- The use of multiple, selected LSM methods can protect a user against false positives because false positives often do not co-occur across methods.

References

- Bates, E. et al. (2003). *Nature Neuroscience*, 6(5), 448–50. doi: 10.1038/nn1050
- Rorden, C. et al. (2007). *J Cognitive Neuroscience*, 19(7), 1081–1088. doi: 10.1162/jocn.2007.19.7.1081
- Mah, Y. et al (2014). *Brain*, 137(9), 2522–2531. doi: 10.1093/brain/awu164
- Pustina, D. et al. (2018). *Neuropsychologia*, 115, 154–166. doi:10.1016/j.neuropsychologia.2017.08.027
- DeMarco, A. & Turkeltaub, P. *Hum Brain Mapp.* 2018 Nov;39(11):4169-4182. doi: 10.1002/hbm.24289
- Wilson, S. et al, *Brain*, 2010 Jul;133(Pt 7):2069-88. doi: 10.1093/brain/awq129

For information/reprints, email - tjherron@ebire.org

D110, Mon AM

