NAVIGATING THE TURBELENT SEAS OF LESION SYMPTOM MAPPING: COMPARATIVE ANALYSIS OF UNIVARIATE AND MULTIVARIATE LESION SYMPTOM MAPPING METHODS

Ivanova, M.V.[1,2,3], Herron, T.[3], Curran, B.[3], Dronkers, N.F.[1,3,4], & Baldo, J.V.[3]

[1] UC Berkeley; [2] National Research University HSE, Russia; [3] VA Northern California Health Care System, CA [4] UC Davis

INTRODUCTION & AIMS OF THE CURRENT STUDY

Lesion symptom mapping (LSM) tools are used to identify brain regions critical for a given behavior.

- **Univariate lesion-symptom mapping (ULSM)** methods provide statistical comparisons of behavioral test scores in patients with and without a lesion on a voxel by voxel basis.
- **Multivariate lesion-symptom mapping (MLSM)** methods consider the effects of all lesioned voxels in one model simultaneously and analyze their contribution to behavior.
- Very little systematic work has been done to empirically outline advantages and disadvantages of these methods.

In the current study we conducted a comprehensive comparison between ULSM and MLSM methods by analyzing their performance under varying conditions.

- Using artificial behavioral data investigated single / dual (network) / zero (pure false positive) anatomical target simulations.
- Explored influence of various factors: anatomical target location, sample size, behavioral noise level, and lesion smoothing.
- Investigated mapping power and spatial accuracy.

METHODS: LSM methods evaluated

Univariate LSM *

- **T-max** Maximum t-value
- **T-nu=125** 125th highest t-value (Mirman et al., 2018)
- **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 ULSM methods used linear regression at every voxel plus permutation testing to set familywise (non-parametric FWER) thresholds based on five different criteria listed above.

Multivariate LSM **

- **SVR** Support vector regression
- **PLS** Partial least squares (dense)
- **ICA-L1**, **ICA-L2** ICA - Independent component analysis
- **LPCA-L1**, **LPCA-L2** LPCA – Logistic principal component analysis
- **SVD-L1**, **SVD-L2** SVD – Singular value decomposition

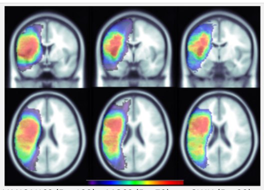
** L1 – elastic net regression; 95% L1 penalty; L2 – elastic net regression; 95% L2 penalty

METHODS: Simulation procedures

Lesion masks from 404 left hemisphere stroke patients:

- Our own database at the VA Northern California Health Care System (n = 209),
- Moss Rehabilitation dataset (n=131) distributed with the LESYMAP software (Pustina et al., 2018);
- George Washington University dataset (n=64) distributed with the SVR software (DeMarco & Turkeltaub, 2019).

For each simulation analysis, the specified number of lesion masks were randomly selected from one of the three datasets (without mixing them together).



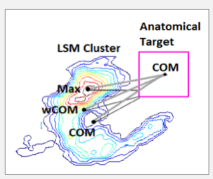
VANCAHCS (5 – 133) MOSS (5 – 70) GWU (5 – 38)

Artificial behavioral scores were based on lesion load to atlas-based anatomical ROIs:

- 16 larger or 30 smaller anatomical ROIs
  - Based on grey matter areas in the left middle cerebral artery region from FSL's version of the Harvard-Oxford atlas and thresholded at 50%.
  - Used 16 such parcels that had 5% or greater area within at least 25% of the lesion masks.
  - To create a set of smaller parcels, each of these 16 parcels was divided into two sections along the axis of maximal spatial extent.

Other parameters explored:

- sample size: n = 32,48,64,80,96,112, 128, & 208;
- behavioral noise level: 0, 0.36, or 0.71 SD of normalized behavioral scores;
- lesion smoothing: 0 mm or 4 mm Gaussian FWHM.



Distance-based measures: grey for COMtarget distances, black for Closesttarget distances.

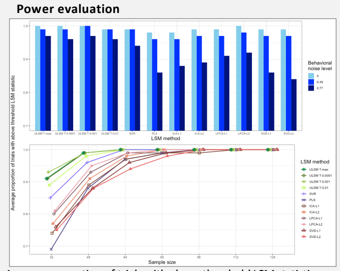
Evaluation measures:

- Power: proportion of trials that yielded any significant LSM statistical values;
- Spatial accuracy:
  - Distance-based (for single target only): mean centroid location (COM), mean centroid location weighted by statistical values (wCOM) & maximum statistic location (Max) of the LSM output map;
  - Overlap-based: dice coefficient & *one-sided Kuiper (OSK) distribution difference;*
- False-positive effects: proportion of trials that yielded above threshold LSM statistic (non-desirable outcome in this instance), and the number and the size of the false positive clusters produced.

We varied these factors in a fully crossed manner in order to systematically compare effect sizes and significance across the different ULSM and MLSM methods for single / dual (network) / zero (pure false positive) anatomical target simulations.

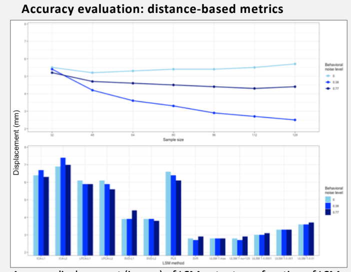
RESULTS: Single anatomical target simulations

Power evaluation



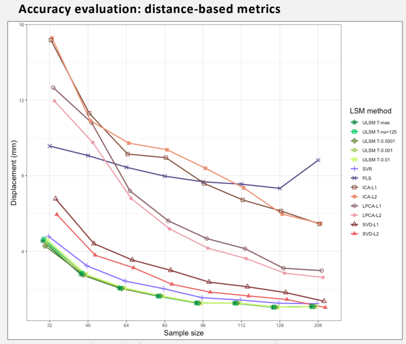
Average proportion of trials with above threshold LSM statistic as a function of LSM method, behavioral noise level and sample size.

Accuracy evaluation: distance-based metrics



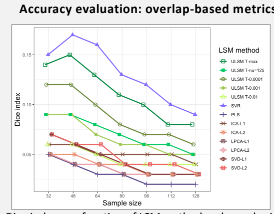
Average displacement (in mm) of LSM output as a function of LSM method, sample size and behavioral noise level.

Accuracy evaluation: distance-based metrics

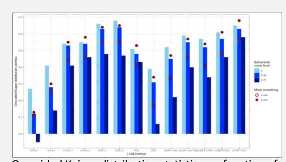


Displacement (in mm) of LSM output map position for single target simulations across different LSM methods at different sample sizes calculated as the average distance between maximum statistic location (Max) and nearest location on the target parcel to the LSM output map.

Accuracy evaluation: overlap-based metrics



Dice index as a function of LSM method and sample size.



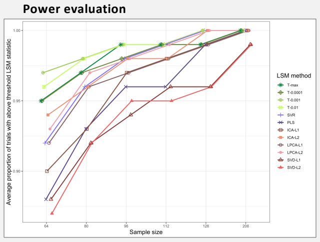
One-sided Kuiper distribution statistic as a function of LSM method, behavioral noise level and mask smoothing.

RESULTS: Dual (network) anatomical target simulations

Three types of networks considered:

- *Redundant* – minimum lesion load of the two target parcels is used to generate the synthetic behavioral score;
- *Extended spatially single-target* – average lesion load of the two parcels;
- *Fragile* – maximal lesion load of the two parcels.

Power evaluation

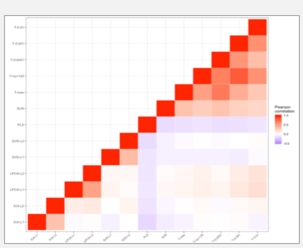


Accuracy evaluation: one-sided Kuiper distribution statistic

LSM method	Redundant	Extended	Fragile
T-max	−0.35	−0.1	−0.17
T-0.0001	−0.23	−0.02	−0.11
T-0.001	−0.1	0.1	0
T-0.01	0.04	0.23	0.13
T-nu=125	−0.19	0.06	−0.02
SVR	−0.43	−0.27	−0.34
PLS	−0.02	0.06	0.03
ICA-L1	−0.38	−0.07	−0.16
ICA-L2	−0.17	0.12	0.04
LPCA-L1	0.07	0.31	0.23
LPCA-L2	0.13	0.34	0.27
SVD-L1	0.11	0.37	0.29
SVD-L2	0.07	0.39	0.3

RESULTS: Zero (false positive) anatomical targets simulations

LSM method	# of Clusters	# of Voxels
T-max	1.5	17
T-nu=125	4.5	312
T-0.0001	1.2	73
T-0.001	1.0	452
T-0.01	1.0	2323
SVR	1.5	17
PLS	5.8	2435
ICA-L1	4.4	715
ICA-L2	4.6	719
LPCA-L1	5.6	1619
LPCA-L2	5.3	1876
SVD-L1	6.9	873
SVD-L2	7.3	963



Correlations between false positive rates of different LSM methods.

DISCUSSION

Single anatomical target simulations demonstrated:

- Good spatial accuracy for ULSM methods with conservative FWER thresholds and some of the simpler DR (e.g., SVD-based) and regression-based (e.g., SVR) MLSM methods;
- Variable accuracy across spatial locations, with especially poor performance in cortical locations on the edge of the lesion masks (areas of lower power);
- More accurate localization with lesion mask smoothing for all LSM methods;
- The importance of having a sample with ≥ 64 patients (with the majority of MLSM methods requiring on average 10-20 more patients to achieve a ULSM level of spatial accuracy);
- Robustness of the maximum statistic as a measure of LSM statistical map location.

Dual anatomical target simulations showed:

- More accurate localization with some of the DR MLSM techniques (e.g., LPCA) as well as ULSM methods with relatively liberal cluster-based FWER thresholds;
- The importance of having a sample with at least ≥ 100 patients.

False positive simulations revealed:

- Cluster sizes were generally the lowest for ULSM methods with conservative FWER thresholds and regression-based MLSM methods.

CONCLUSIONS

- Our simulations show no clear superiority of MLSM techniques over the ULSM methods.
- Depending on the design of a particular LSM study and specific hypothesis regarding the expected brain-behavior relationship, different LSM methods are indicated.
- It is advantageous to implement both ULSM and MLSM methods in tandem to enhance confidence in the results, as significant matching foci identified with both methods are unlikely to be spurious.

For information/reprints: ivanova@berkeley.edu