METHODS: Simulation procedures

Lesion masks from 404 left hemisphere stroke patients:
- Our own database at the VA Northern California Health Care System (n = 209).
- Mose Rehabilitation database (n = 131) distributed with the LESTMAP software (Ponzi et al., 2018).
- George Washington University dataset (n = 64) distributed with the SVR software (DeMarco & Turltehaard, 2019).

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

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 = 52, 46, 84, 90, 96, 112, 128, & 208.
- Behavioral noise level: 0.3; 0.5; 0.7 7 (0.01 SD of normalized behavioral scores; lesion smoothing: 0 mm or 4 mm Gaussian FWHM.

Evaluation measures:
- Power: proportion of trials that yielded any significant ULSM statistical values;
- Spatial accuracy: distance (for single target only); mean centered location (LRU) coordinate distribution;
- Overlap-based: dice coefficient & one-sided kiper (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 size of the false positive clusters produced.

RESULTS: Dual (network) anatomical target simulations

Three types of networks considered:
- Redundant – maximum 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 – maximum lesion load of the two parcels.

DISCUSSION

Single anatomical target simulations demonstrated:
- Good spatial accuracy for ULSM methods with conservative FWER thresholds and some of the simpler DR (e.g., SVR) and regression-based (e.g., SVR) ULSM methods;
- Variable accuracy across spatial locations, with especially poor performance in cortical locations on the edge of the lesion mask (areas of lower power);
- More accurate localization with lesion mask smoothing for all ULSM methods;
- The importance of having a sample of ≥ 64 patients (with the majority of ULSM 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 ULSM 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.

CONCLUSIONS

- Our simulations show no clear superiority of MLSM techniques over the ULSM methods.
- Depending on the design of a particular ULSM study and specific hypothesis regarding the expected brain-behavior relationship, different ULSM 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.